(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 May 2001 (03.05.2001)

PCT

(10) International Publication Number
WO 01/30964 A2

(51) International Patent Classification[7]: C12N

(21) International Application Number: PCT/US00/29126

(22) International Filing Date: 20 October 2000 (20.10.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/161,232	22 October 1999 (22.10.1999)	US
Not furnished	19 October 2000 (19.10.2000)	US

(71) Applicant *(for all designated States except US)*: **LIFESPAN BIOSCIENCES, INC.** [US/US]; 700 Blanchard Street, Seattle, WA 98121 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **BURMER, Glenna, C.** [US/US]; 7516 55th Place, N.E., Seattle, WA 98115 (US). **BROWN, Joseph, P.** [GB/US]; 411 West Prospect Street, Seattle, WA 98119 (US). **PRITCHARD, David** [GB/US]; 3429 Burke Avenue, N #34, Seattle, WA 98103 (US).

(74) Agents: **PARENT, Annette, S.** et al.; Townsend and Townsend and Crew LLP, Two Embarcadero Center, Eighth Floor, San Francisco, CA 94111 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/30964 A2

(54) Title: ANTI-CANCER NUCLEIC ACID AND PROTEIN TARGETS

(57) Abstract: This invention relates to the discovery of nucleic acids associated with cell proliferation, neoplasia, cell transformation, malignant tumor formation and metastasis and uses therefor.

ANTI-CANCER NUCLEIC ACID AND PROTEIN TARGETS

CROSS-REFERENCE TO RELATED APPLICATION

The present application claims priority to U.S. Patent Application No. 60/161,232, filed October 22, 1999, herein incorporated by reference in its entirety.

FIELD OF THE INVENTION

This invention relates to the discovery of nucleic acids and proteins associated with cancer. The identification of these cancer-associated nucleic acids and proteins have diagnostic uses in detecting the cancerous state of a cell population as well as applications for gene therapy and the prevention of malignant tumor development.

BACKGROUND OF THE INVENTION

Despite years of research into its causes and potential treatments, cancer remains the second leading cause of death in the United States. Currently, more than 500 thousand Americans die of cancer each year, more than 1500 people a day. In addition, even more cases will be detected: in 1999, more than 1.2 million new cancer cases will be diagnosed. Of the many types of cancer, epithelial cancers are among the most prevalent and deadly. For example, about 175,000 new invasive cases of breast cancer are expected to occur among women in the United States during 1999, and more than 43,000 women will die of this disease. Although some progress has been made towards understanding the causes of various types of cancer, a major need remains for new tools for the diagnosis and treatment of this disease.

Cancer is a genetic disease of single cell origin caused by the accumulation of inherited and acquired mutations in specific cancer-associated genes, which have normal cellular functions, but when mutated or present at abnormally high levels contribute to cancer. Such cancer-associated genes can be

classified in two categories: genes that cause cancer when they become inactivated and genes that lead to cancer when they are mutated or overexpressed (Alberts *et al.* (1994) *Molecular Biology of the Cell*, Third Edition. Garland Publishing, New-York, pp: 1255-1291). The first group of genes are often referred to as tumor-suppressor genes and include, *e.g.*, the retinoblastoma gene, p53, *etc.* (Holden *et al.* (1999) *Med. Hypotheses* 52:483-485; McCormick, F. (1999) *Cancer J. Sci. Am.* 5:139-144; Prives *et al.*, (1999) *J. Pathol.* 187:112-126; Vooijs et al. (1999) *Oncogene* 18:5293-5303; Stiegler *et al.* (1998) *J. Cell Biochem.* Suppl. 30:30-36). They are required for the normal functioning of a cell, for example by restricting cellular proliferation or growth, or by controlling cell adhesion. Their inactivation or loss of their function results in the loss of these controls or properties and contributes to tumor development (Dictor *et al.* (1999) *Am. J. Clin. Pathol.* 112:S40-52; Bartek et al. (1997) *Exp. Cell Res.* 237:1-6; Kaelin (1997) *Ann. N. Y. Acad. Sci.* 833:29-33). The genes belonging to the second class are referred to as oncogenes and they include, *e.g.*, ras, myc, src, *etc.* (Introna *et al.* (1999) *Leukemia* 13: 1301-1306; Nesbit *et al.* (1999) *Oncogene* 18:3004-3016; Schwab (1999) *Semin. Cancer Biol.* 9:319-125; Stice *et al.* (1999) *Front Biosci.* 15:D72-86; Porter *et al.* (1998) *Oncogene* 17:1343-1352). A mutation or any other event resulting in their overexpression can also contribute to cancer development, for example, by releasing cells from normal restraints on proliferation. Although several genes have been found in each of these categories, the mechanism underlying cancer development is still largely unknown. In addition, it is well established that cancer is often the result of a number of mutations in different genes (Alberts *et al.* (1994) *Molecular Biology of the Cell*, Third Edition. Garland Publishing, New-York, pp: 1255-1291; Devereux *et al.* (1999) *IARC Sci. Publ.* 146: 19-42; Fearon *et al.* (1999) *Curr. Biol.* 28:R62-65; Murakami et al. (1998) *Mutat. Res.* 400: 421-437). There is therefore a major need in the field for identifying both genes that control the normal functioning of a cell and prevent its transformation into a malignant cell and genes whose expression needs to be strictly controlled to prevent cancer development. This knowledge would provide a major tool for diagnosing and treating cancer.

The present invention is based on the discovery that a number of polynucleotide sequences or proteins are overexpressed in many types of cancer cells in mammals, while other polynucleotide sequences or proteins are

underexpressed in cancerous cells, compared to normal healthy tissues. As described herein, this discovery has provided novel and badly needed diagnostic, prognostic, and therapeutic tools for many types of cancers.

SUMMARY OF THE INVENTION

The present invention provides isolated nucleic acids and proteins associated with cancer. Such sequences can be used to determine the cancerous state of a cell population, *e.g.*, whether cells are proliferating at an abnormally high rate, are undergoing transformation and/or are malignant. Moreover, the present invention provides sequences indicative of the proliferation state, transformation, malignancy or metastatic ability of a cell. The genes that exhibit cancer-associated alterations in expression can also be targeted and their level of expression altered by, for example, gene therapy methods (*e.g.*, by altering the subject sequences). Such methods can be used, for example, to slow or stop the tumor development process, to arrest proliferation of a cell population, such as a tumor cell population, to inhibit transformation of a cell and to determine that a cell population is healthy.

In one aspect, the present invention provides a method for diagnosing cancer of a tissue of interest in a patient, comprising: detecting the overexpression or the underexpression of a cancer-associated molecule in the tissue of interest in the patient, where overexpression or underexpression of such molecule is indicative of a cancer. In a preferred embodiment, the cancer-associated molecule is selected from the molecules listed in Table 1. In one embodiment, underexpression of the cancer-associated molecule is indicative of cancer, and the cancer-associated molecule is underexpressed in the patient. In another embodiment, overexpression of the cancer-associated molecule is indicative of cancer, and the cancer-associated molecule is overexpressed in the patient. In a preferred embodiment, the cancer-associated molecule is an mRNA. In another preferred embodiment, the cancer-associated molecule is detected using an immunoassay. The tissue of interest is preferably selected from the group consisting of liver, breast, prostate, kidney and colon.

In another aspect, the invention provides a method for arresting cancer, comprising introducing into a cell a cancer-associated molecule according to

Table 1, wherein underexpression of the cancer-associated molecule is indicative of cancer. The cancer-associated molecule of interest may be a nucleic acid encoding a cancer-associated protein or a protein.

In still another aspect, the invention provides a method for arresting cancer comprising inhibiting a cancer-associated molecule according to Table 1, wherein overexpression of the cancer-associated molecule is indicative of cancer. In a preferred embodiment, the cancer-associated molecule is inhibited using an antisense polynucleotide. In another preferred embodiment, the cancer-associated molecule is inhibited using an antibody that specifically binds to said cancer-associated protein.

The present invention also provides a method for identifying a modulator of cancer development, the method comprising: culturing a cell in the presence of the modulator to form a first cell culture; contacting RNA or cDNA from the first cell culture with a probe which comprises a polynucleotide sequence associated with cancer; determining whether the amount of the probe which hybridizes to the RNA or cDNA from the first cell culture is increased or decreased relative to the amount of the probe which hybridizes to RNA or cDNA from a second cell culture grown in the absence of the modulator, and further, detecting a phenotype indicative of an amelioration of the cancerous state of the cell population that is treated with the modulator. In a preferred embodiment, the probe comprises at least about 10 nucleotides from a polynucleotide sequence selected from the group consisting of the sequences set out in Table 1 which show increased expression in cancerous tissues. Amelioration of the cancerous state of a cell population may include, for example, a decrease in the proliferative potential, transformation, malignancy and/or metastatic ability of a cell. The first and second cell cultures may be obtained, for example, from a liver, a kidney, a breast, a colon or a prostate cell.

In still another aspect, the present invention provides kits for carrying out the various methods. For instance, in one embodiment, a kit is provided for detecting whether a cell is undergoing transformation or becoming malignant, the kit comprising: a probe which comprises a polynucleotide sequence associated with cancer and a label for detecting the presence of the probe. In one embodiment, the probe comprises at least about 10 nucleotides from a polynucleotide sequence

selected from the group consisting of the sequences listed in Table 1. Additionally, this kit can further comprise a plurality of probes each of which comprises a polynucleotide sequence associated with cancer; and a label or labels for detecting the presence of the plurality of probes. The probes can optionally be immobilized on a solid support (*e.g.*, a chip).

The invention also embraces the use of antisense methods for examining cancer in animals and cells. Typically, any time a gene is identified, it can be examined by knocking out the gene in an animal and observing the effect on the animal phenotype. Knockouts can be achieved by transposons which insert by homologous recombinations, antisense or ribozymes specifically directed at disturbing the embryonic stem cells of an organism such as a mouse. Ribozymes can include any of the various types of ribozymes modified to cleave the mRNA encoding, for example, the cancer-associated protein. Examples include hairpins and hammerhead ribozymes. Finally, antisense molecules which selectively bind, for example, to the mRNA encoding a cancer-associated protein are expressed via expression cassettes operably linked to subsequences of the cancer-associated gene and generally comprise 20-50 base long sequences in opposite orientation to the mRNA to which they are targeted.

The invention also provides a method for modulating cancer development in a patient in need thereof, the method comprising administering to the patient a compound that modulates cancer development. In a preferred embodiment, the modulator increases or decreases the expression of a nucleic acid sequence associated with cancerous transformation of a cell, *e.g.*, a liver, kidney, breast, colon or prostate cell, set out in Table 1. In this embodiment, the nucleic acid can, for example, comprise at least about 10 nucleotides from a polynucleotide sequence selected from the group consisting of the sequences set out in Table 1.

DEFINITIONS

"Cancer" or "malignancy" are used as synonymous terms and refer to any of a number of diseases that are characterized by uncontrolled, abnormal proliferation of cells, the ability of affected cells to spread locally or through the

bloodstream and lymphatic system to other parts of the body (i.e. metastasize) as well as any of a number of characteristic structural and/or molecular features. A "cancerous" or "malignant cell" is understood as a cell having specific structural properties, lacking differentiation and being capable of invasion and metastasis. Examples of cancers are kidney, colon, breast, prostate and liver cancer.

"Neoplasia" or "tumor" describe new abnormal growth which can be benign or malignant. "Neoplastic proliferation" describes cell proliferation that persists even in the absence of growth stimulus.

In the context of the invention, the term "transformation" refers to the change that a normal cell undergoes as it becomes malignant. In eukaryotes, the term "transformation" can be used to describe the conversion of normal cells to malignant cells in cell culture.

"Proliferating cells" are those which are actively undergoing cell division and growing exponentially. "Loss of cell proliferation control" refers to the property of cells that have lost the cell cycle controls that normally ensure appropriate restriction of cell division. Cells that have lost such controls proliferate at a faster than normal rate, without stimulatory signals, and do not respond to inhibitory signals.

"Amplification" primers are oligonucleotides comprising either natural or analog nucleotides that can serve as the basis for the amplification of a select nucleic acid sequence. They include, for example, both polymerase chain reaction primers and ligase chain reaction oligonucleotides.

"Antibody" refers to a polypeptide substantially encoded by an immunoglobulin gene or immunoglobulin genes, or fragments thereof which specifically bind and recognize an analyte (antigen). The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively.

An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one "light" (about 25 kD) and one "heavy" chain (about 50-70 kD).

The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily responsible for antigen recognition. The terms variable light chain (V_L) and variable heavy chain (V_H) refer to these light and heavy chains respectively.

Antibodies exist, *e.g.*, as intact immunoglobulins or as a number of well characterized fragments produced by digestion with various peptidases. Thus, for example, pepsin digests an antibody below the disulfide linkages in the hinge region to produce F(ab)'$_2$, a dimer of Fab which itself is a light chain joined to V_H-C_H1 by a disulfide bond. The F(ab)'$_2$ may be reduced under mild conditions to break the disulfide linkage in the hinge region, thereby converting the F(ab)'$_2$ dimer into an Fab' monomer. The Fab' monomer is essentially an Fab with part of the hinge region (*see*, *Fundamental Immunology*, Third Edition, W.E. Paul, ed., Raven Press, N.Y. 1993). While various antibody fragments are defined in terms of the digestion of an intact antibody, one of skill will appreciate that such fragments may be synthesized *de novo* either chemically or by utilizing recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments either produced by the modification of whole antibodies or those synthesized *de novo* using recombinant DNA methodologies (*e.g.*, single chain Fv).

"Cancer-associated" refers to the relationship of a nucleic acids and its expression, or lack thereof, or a protein and its level or activity, or lack thereof, to the onset of malignancy in a subject cell. For example, cancer can be associated with expression of a particular gene that is not expressed, or is expressed at a lower level, in a normal healthy cell. Conversely, a cancer-associated gene can be one that is not expressed in a malignant cell (or in a cell undergoing transformation), or is expressed at a lower level in the malignant cell than it is expressed in a normal healthy cell.

"Biological samples" refers to any tissue or liquid sample having genomic DNA or other nucleic acids (*e.g.*, mRNA) or proteins. It refers to samples of cells with a normal complement of chromosomes as well as samples of cells suspected of malignancy.

The term "gene" means the segment of DNA involved in producing a polypeptide chain; it includes regions preceding and following the coding region

(leader and trailer) as well as intervening sequences (introns) between individual coding segments (exons).

The term "isolated," when applied to a nucleic acid or protein, denotes that the nucleic acid or protein is essentially free of other cellular components with which it is associated in the natural state. It is preferably in a homogeneous state although it can be in either a dry or aqueous solution. Purity and homogeneity are typically determined using analytical chemistry techniques such as polyacrylamide gel electrophoresis or high performance liquid chromatography. A protein which is the predominant species present in a preparation is substantially purified. In particular, an isolated gene is separated from open reading frames which flank the gene and encode a protein other than the gene of interest. The term "purified" denotes that a nucleic acid or protein gives rise to essentially one band in an electrophoretic gel. Particularly, it means that the nucleic acid or protein is at least 85% pure, more preferably at least 95% pure, and most preferably at least 99% pure.

The term "nucleic acid" refers to deoxyribonucleotides or ribonucleotides and polymers thereof in either single- or double-stranded form. Unless specifically limited, the term encompasses nucleic acids containing known analogues of natural nucleotides which have similar binding properties as the reference nucleic acid and are metabolized in a manner similar to naturally occurring nucleotides. Unless otherwise indicated, a particular nucleic acid sequence also implicitly encompasses conservatively modified variants thereof (*e.g.*, degenerate codon substitutions) and complementary sequences and as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues (Batzer *et al.*, *Nucleic Acid Res. 19*:5081 (1991); Ohtsuka *et al.*, *J. Biol. Chem. 260:*2605-2608 (1985); and Cassol *et al.*, 1992; Rossolini *et al.*, *Mol. Cell. Probes 8*:91-98 (1994)). The term nucleic acid is used interchangeably with gene, cDNA, and mRNA encoded by a gene.

As used herein a "nucleic acid probe" is defined as a nucleic acid capable of binding to a target nucleic acid (*e.g.*, a nucleic acid associated with cancer) of complementary sequence through one or more types of chemical bonds,

usually through complementary base pairing, usually through hydrogen bond formation. As used herein, a probe may include natural (*i.e.,* A, G, C, or T) or modified bases (7-deazaguanosine, inosine, *etc.*). In addition, the bases in a probe may be joined by a linkage other than a phosphodiester bond, so long as it does not interfere with hybridization. Thus, for example, probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages. It will be understood by one of skill in the art that probes may bind target sequences lacking complete complementarity with the probe sequence depending upon the stringency of the hybridization conditions.

Nucleic acid probes can be DNA or RNA fragments. DNA fragments can be prepared, for example, by digesting plasmid DNA, or by use of PCR, or synthesized by either the phosphoramidite method described by Beaucage and Carruthers, *Tetrahedron Lett.* 22:1859-1862 (1981) (Beaucage and Carruthers), or by the triester method according to Matteucci, *et al., J. Am. Chem. Soc.,* 103:3185 (1981) (Matteucci), both incorporated herein by reference. A double stranded fragment may then be obtained, if desired, by annealing the chemically synthesized single strands together under appropriate conditions, or by synthesizing the complementary strand using DNA polymerase with an appropriate primer sequence. Where a specific sequence for a nucleic acid probe is given, it is understood that the complementary strand is also identified and included. The complementary strand will work equally well in situations where the target is a double-stranded nucleic acid.

A "labeled nucleic acid probe" is a nucleic acid probe that is bound, either covalently, through a linker, or through ionic, van der Waals or hydrogen bonds to a label such that the presence of the probe may be detected by detecting the presence of the label bound to the probe.

The phrase "a nucleic acid sequence encoding" refers to a nucleic acid which contains sequence information for a structural RNA such as rRNA, a tRNA, or the primary amino acid sequence of a specific protein or peptide, or a binding site for a trans-acting regulatory agent. This phrase specifically encompasses degenerate codons (*i.e.,* different codons which encode a single amino acid) of the native sequence or sequences which may be introduced to conform with codon preference in a specific host cell.

"Stringent hybridization conditions" and "stringent hybridization wash conditions" in the context of nucleic acid hybridization experiments, such as Southern and northern hybridizations, are sequence dependent, and are different under different environmental parameters. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes*, part I, chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays," Elsevier, NY. Generally, highly stringent hybridization and wash conditions are selected to be about 5° C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. Typically, under "stringent conditions," a probe will hybridize to its target subsequence, but to no other sequences.

The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Very stringent conditions are selected to be equal to the T_m for a particular probe. An example of stringent hybridization conditions for hybridization of complementary nucleic acids which have more than 100 complementary residues on a filter in a Southern or northern blot is 50% formamide with 1 mg of heparin at 42°C, with the hybridization being carried out overnight. An example of highly stringent wash conditions is 0.15M NaCl at 72°C for about 15 minutes. An example of stringent wash conditions is a 0.2x SSC wash at 65°C for 15 minutes (*see,* Sambrook, *supra*, for a description of SSC buffer). Often, a high stringency wash is preceded by a low stringency wash to remove background probe signal. An example medium stringency wash for a duplex of, *e.g.,* more than 100 nucleotides, is 1x SSC at 45°C for 15 minutes. An example low stringency wash for a duplex of, *e.g.,* more than 100 nucleotides, is 4-6x SSC at 40°C for 15 minutes. For short probes (*e.g.,* about 10 to 50 nucleotides), stringent conditions typically involve salt concentrations of less than about 1.0 M Na ion, typically about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3, and the temperature is typically at least about 30°C. Stringent conditions can also be achieved with the addition of destabilizing agents such as formamide. In general, a signal to noise ratio of 2x (or higher) than that observed for an unrelated probe in the particular hybridization assay indicates detection of a specific hybridization. Nucleic acids which do not hybridize to each

other under stringent conditions are still substantially identical if the polypeptides which they encode are substantially identical. This occurs, *e.g.*, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code.

The phrase "specifically (or selectively) binds to an antibody" or "specifically (or selectively) immunoreactive with", when referring to a protein or peptide, refers to a binding reaction which is determinative of the presence of the protein in the presence of a heterogeneous population of proteins and other biologics. Thus, under designated immunoassay conditions, the specified antibodies bind to a particular protein and do not bind in a significant amount to other proteins present in the sample. Specific binding to an antibody under such conditions may require an antibody that is selected for its specificity for a particular protein. For example, antibodies raised to the protein with the amino acid sequence encoded by any of the polynucleotides of the invention can be selected to obtain antibodies specifically immunoreactive with that protein and not with other proteins except for polymorphic variants. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA immunoassays, Western blots, or immunohistochemistry are routinely used to select monoclonal antibodies specifically immunoreactive with a protein. *See*, Harlow and Lane (1988) *Antibodies, A Laboratory Manual*, Cold Spring Harbor Publications, New York ("Harlow and Lane") for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity. Typically, a specific or selective reaction will be at least twice background signal or noise and more typically more than 10 to 100 times background.

DETAILED DESCRIPTION AND PREFERRED EMBODIMENTS OF THE INVENTION

The present invention provides nucleic acids and proteins that are indicative of cancer. Host cells, vectors, and probes are described, as are antibodies to the proteins and uses of the proteins as antigens. The present invention provides methods for obtaining and expressing nucleic acids, methods for purifying gene

products, other methods that can be used to detect and quantify the expression and quality of the gene product (*e.g.*, proteins), and uses for both the nucleic acids and the gene products.

This invention relies on routine techniques in the field of recombinant genetics. A basic text disclosing the general methods of use in this invention is Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Publish., Cold Spring Harbor, NY 2nd ed. (1989) and Kriegler, *Gene Transfer and Expression: A Laboratory Manual*, W.H. Freeman, N.Y., (1990), which are both incorporated herein by reference. Unless otherwise stated all enzymes are used in accordance with the manufacturer's instructions.

I. Detection of Gene Expression and Genomic Analysis of Cancer-Associated Proteins.

The polynucleotides and polypeptides of the present invention can be employed as research reagents and materials for discovery of treatments and diagnostics to human disease. It will be readily apparent to those of skill in the art that although the following discussion is directed to methods for detecting nucleic acids associated with cancer, similar methods can be used to detect nucleic acids associated with, *e.g.*, cell proliferation, cell transformation, neoplasia, metastasis, *etc.*

As should be apparent to those of skill in the art, the invention is the identification of cancer-associated genes and the discovery that multiple nucleic acids are associated with cancer development, cell proliferation, cell transformation, neoplasia and/or metastasis. Accordingly, the present invention also includes methods for detecting the presence, alteration or absence of the such cancer-associated nucleic acids (*e.g.*, DNA or RNA) in a physiological specimen in order to determine the level of proliferation and/or transformation of cells *in vitro*, or *ex vivo*, as well as the genotype and risk of cancer development associated with mutations created in cancer-associated sequences. Although any tissue having cells bearing the genome of an individual, or RNA associated with cancer, can be used, the most convenient specimen will be blood samples or biopsies of suspect tissue. It is also possible and preferred in some circumstances to conduct assays on cells that are

isolated under microscopic visualization. A particularly useful method is the microdissection technique described in PCT Published Application No. WO 95/23960. The cells isolated by microscopic visualization can be used in any of the assays described herein including both genomic and immunological based assays.

This invention provides for methods of genotyping family members in which relatives are diagnosed with cancer. Conventional methods of genotyping are provided herein.

The invention provides methods for detecting whether a cell is in a cancerous state or is undergoing transformation and/or becoming malignant. The methods typically comprise contacting RNA from the cell with a probe which comprises a polynucleotide sequence associated with cancer, and determining whether the amount of the probe which hybridizes to the RNA is increased or decreased relative to the amount of the probe which hybridizes to RNA from a non-cancerous cell. The assays are useful for detecting cell transformation associated with, for example, different types of cancer. One can also detect cell proliferation, neoplasia and/or transformation using the methods of the invention.

The probes are capable of binding to a target nucleic acid (*e.g.*, a nucleic acid associated with cancer development). By assaying for the presence or absence of the probe, one can detect the presence or absence of the target nucleic acid in a sample. Preferably, non-hybridizing probe and target nucleic acids are removed (*e.g.*, by washing) prior to detecting the presence of the probe.

A variety of methods of specific DNA and RNA measurement using nucleic acid hybridization techniques are known to those of skill in the art. *See*, Sambrook, *supra*. For example, one method for evaluating the presence or absence of the DNA in a sample involves a Southern transfer. Briefly, the digested genomic DNA is run on agarose slab gels in buffer and transferred to membranes. Hybridization is carried out using the probes discussed above. Visualization of the hybridized portions allows the qualitative determination of the presence, alteration or absence of a cancer-associated gene.

Similarly, a Northern transfer may be used for the detection of cancer-associated mRNA in samples of RNA from cells expressing the cancer-associated proteins. In brief, the mRNA is isolated from a given cell sample using an acid guanidinium-phenol-chloroform extraction method. The mRNA is then

electrophoresed to separate the mRNA species and the mRNA is transferred from the gel to a nitrocellulose membrane. As with the Southern blots, labeled probes are used to identify the presence or absence of the subject protein transcript. Alternatively, the amount of, for example, a cancer-associated mRNA can be analyzed in the absence of electrophoretic separation.

The selection of a nucleic acid hybridization format is not critical. A variety of nucleic acid hybridization formats are known to those skilled in the art. For example, common formats include sandwich assays and competition or displacement assays. Hybridization techniques are generally described in "*Nucleic Acid Hybridization, A Practical Approach*," Ed. Hames, B.D. and Higgins, S.J., IRL Press, 1985; Gall and Pardue (1969), *Proc. Natl. Acad. Sci.*, U.S.A., 63:378-383; and John, Burnsteil and Jones (1969) *Nature*, 223:582-587.

For example, sandwich assays are commercially useful hybridization assays for detecting or isolating nucleic acids. Such assays utilize a "capture" nucleic acid covalently immobilized to a solid support and labeled "signal" nucleic acid in solution. The clinical sample will provide the target nucleic acid. The "capture" nucleic acid and "signal" nucleic acid probe hybridize with the target nucleic acid to form a "sandwich" hybridization complex. To be effective, the signal nucleic acid cannot hybridize with the capture nucleic acid.

Detection of a hybridization complex may require the binding of a signal generating complex to a duplex of target and probe polynucleotides or nucleic acids. Typically, such binding occurs through ligand and anti-ligand interactions as between a ligand-conjugated probe and an anti-ligand conjugated with a signal. The binding of the signal generation complex is also readily amenable to accelerations by exposure to ultrasonic energy.

The label may also allow indirect detection of the hybridization complex. For example, where the label is a hapten or antigen, the sample can be detected by using antibodies. In these systems, a signal is generated by attaching fluorescent or enzyme molecules to the antibodies or in some cases, by attachment to a radioactive label (*see, e.g.,* Tijssen, P., "Practice and Theory of Enzyme Immunoassays," *Laboratory Techniques in Biochemistry and Molecular Biology*, Burdon, R.H., van Knippenberg, P.H., Eds., Elsevier (1985), pp. 9-20).

The probes are typically labeled directly, as with isotopes, chromophores, lumiphores, chromogens, or indirectly labeled such as with biotin to which a streptavidin complex may later bind. Thus, the detectable labels used in the assays of the present invention can be primary labels (where the label comprises an element that is detected directly or that produces a directly detectable element) or secondary labels (where the detected label binds to a primary label, *e.g.*, as is common in immunological labeling). Typically, labeled signal nucleic acids are used to detect hybridization. Complementary nucleic acids or signal nucleic acids may be labeled by any one of several methods typically used to detect the presence of hybridized polynucleotides. The most common method of detection is the use of autoradiography with ^{3}H, ^{125}I, ^{35}S, ^{14}C, or ^{32}P-labeled probes or the like.

Other labels include ligands which bind to labeled antibodies, fluorophores, chemi-luminescent agents, enzymes, and antibodies which can serve as specific binding pair members for a labeled ligand. An introduction to labels, labeling procedures and detection of labels is found in Polak and Van Noorden (1997) *Introduction to Immunocytochemistry*, 2nd ed., Springer Verlag, New York, and in Haugland (1996) *Handbook of Fluorescent Probes and Research Chemicals*, a combined handbook and catalogue Published by Molecular Probes, Inc., Eugene, OR. Primary and secondary labels can include undetected elements as well as detected elements. Useful primary and secondary labels in the present invention can include spectral labels such as fluorescent dyes (*e.g.*, fluorescein and derivatives such as fluorescein isothiocyanate (FITC) and Oregon Green™, rhodamine and derivatives (*e.g.*, Texas red, tetrarhodimine isothiocynate (TRITC), *etc.*), digoxigenin, biotin, phycoerythrin, AMCA, CyDyes™, and the like), radiolabels (*e.g.*, ^{3}H, ^{125}I, ^{35}S, ^{14}C, ^{32}P, ^{33}P, *etc.*), enzymes (*e.g.*, horse radish peroxidase, alkaline phosphatase *etc.*), spectral colorimetric labels such as colloidal gold or colored glass or plastic (*e.g.*, polystyrene, polypropylene, latex, *etc.*) beads. The label may be coupled directly or indirectly to a component of the detection assay (*e.g.*, the probe) according to methods well known in the art. As indicated above, a wide variety of labels may be used, with the choice of label depending on sensitivity required, ease of conjugation with the compound, stability requirements, available instrumentation, and disposal provisions.

Preferred labels include those that use: 1) chemiluminescence (using horseradish peroxidase and/or alkaline phosphatase with substrates that produce photons as breakdown products as described above) with kits being available, *e.g.*, from Molecular Probes, Amersham, Boehringer-Mannheim, and Life Technologies/ Gibco BRL; 2) color production (using both horseradish peroxidase and/or alkaline phosphatase with substrates that produce a colored precipitate [kits available from Life Technologies/Gibco BRL, and Boehringer-Mannheim]); 3) hemifluorescence using, *e.g.*, alkaline phosphatase and the substrate AttoPhos [Amersham] or other substrates that produce fluorescent products, 4) fluorescence (*e.g.*, using Cy-5 [Amersham]), fluorescein, and other fluorescent tags]; and 5) radioactivity. Other methods for labeling and detection will be readily apparent to one skilled in the art.

Preferred enzymes that can be conjugated to detection reagents of the invention include, *e.g.*, β-galactosidase, luciferase, horse radish peroxidase, and alkaline phosphatase. The chemiluminescent substrate for luciferase is luciferin. One embodiment of a chemiluminescent substrate for β-galactosidase is 4-methylumbelliferyl-β-D-galactoside. Embodiments of alkaline phosphatase substrates include p-nitrophenyl phosphate (pNPP), which is detected with a spectrophotometer; 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium (BCIP/NBT) and fast red/napthol AS-TR phosphate, which are detected visually; and 4-methoxy-4-(3-phosphonophenyl) spiro[1,2-dioxetane-3,2'-adamantane], which is detected with a luminometer. Embodiments of horse radish peroxidase substrates include 2,2'azino-bis(3-ethylbenzthiazoline-6 sulfonic acid) (ABTS), 5-aminosalicylic acid (5AS), o-dianisidine, and o-phenylenediamine (OPD), which are detected with a spectrophotometer; and 3,3,5,5'-tetramethylbenzidine (TMB), 3,3'diaminobenzidine (DAB), 3-amino-9-ethylcarbazole (AEC), and 4-chloro-1-naphthol (4C1N), which are detected visually. Other suitable substrates are known to those skilled in the art. The enzyme-substrate reaction and product detection are performed according to standard procedures known to those skilled in the art and kits for performing enzyme immunoassays are available as described above.

In general, a detector which monitors a particular probe or probe combination is used to detect the detection reagent label. Typical detectors include spectrophotometers, phototubes and photodiodes, microscopes, scintillation counters, cameras, film and the like, as well as combinations thereof. Examples of

suitable detectors are widely available from a variety of commercial sources known to persons of skill in the art. Commonly, an optical image of a substrate comprising bound labeling moieties is digitized for subsequent computer analysis.

Most typically, the amount of, for example, a cancer-associated RNA is measured by quantitating the amount of label fixed to the solid support by binding of the detection reagent. Typically, presence of a modulator during incubation will increase or decrease the amount of label fixed to the solid support relative to a control incubation which does not comprise the modulator, or as compared to a baseline established for a particular reaction type. Means of detecting and quantitating labels are well known to those of skill in the art. Thus, for example, where the label is a radioactive label, means for detection include a scintillation counter or photographic film as in autoradiography. Where the label is optically detectable, typical detectors include microscopes, cameras, phototubes and photodiodes and many other detection systems which are widely available.

In preferred embodiments, the target nucleic acid or the probe is immobilized on a solid support. Solid supports suitable for use in the assays of the invention are known to those of skill in the art. As used herein, a solid support is a matrix of material in a substantially fixed arrangement. Exemplar solid supports include glasses, plastics, polymers, metals, metalloids, ceramics, organics, *etc.* Solid supports can be flat or planar, or can have substantially different conformations. For example, the substrate can exist as particles, beads, strands, precipitates, gels, sheets, tubing, spheres, containers, capillaries, pads, slices, films, plates, dipsticks, slides, *etc.* Magnetic beads or particles, such as magnetic latex beads and iron oxide particles, are examples of solid substrates that can be used in the methods of the invention. Magnetic particles are described in, for example, US Patent No. 4,672,040, and are commercially available from, for example, PerSeptive Biosystems, Inc. (Framingham MA), Ciba Corning (Medfield MA), Bangs Laboratories (Carmel IN), and BioQuest, Inc. (Atkinson NH). The substrate is chosen to maximize signal to noise ratios, primarily to minimize background binding, for ease of washing and cost.

A variety of automated solid-phase assay techniques are also appropriate. For instance, very large scale immobilized polymer arrays (VLSIPS™), available from Affymetrix, Inc. in Santa Clara, CA can be used to

detect changes in expression levels of a plurality of cancer-associated nucleic acids simultaneously. *See*, Tijssen, *supra.*, Fodor *et al.* (1991) *Science*, 251: 767- 777; Sheldon *et al.* (1993) *Clinical Chemistry* 39(4): 718-719, and Kozal *et al.* (1996) *Nature Medicine* 2(7): 753-759. Thus, in one embodiment, the invention provides methods of detecting expression levels of cancer-associated nucleic acids, in which nucleic acids (*e.g.*, RNA from a cell culture) are hybridized to an array of nucleic acids that are known to be associated with malignant tumor development. For example, in the assay described, *supra*, oligonucleotides which hybridize to a plurality of cancer-associated nucleic acids are optionally synthesized on a DNA chip (such chips are available from Affymetrix) and the RNA from a biological sample, such as a cell culture, is hybridized to the chip for simultaneous analysis of multiple cancer-associated nucleic acids. The cancer-associated nucleic acids that are present in the sample which is assayed are detected at specific positions on the chip.

Detection can be accomplished, for example, by using a labeled detection moiety that binds specifically to duplex nucleic acids (*e.g.*, an antibody that is specific for RNA-DNA duplexes). One preferred example uses an antibody that recognizes DNA-RNA heteroduplexes in which the antibody is linked to an enzyme (typically by recombinant or covalent chemical bonding). The antibody is detected when the enzyme reacts with its substrate, producing a detectable product. Coutlee *et al.* (1989) *Analytical Biochemistry* 181:153-162; Bogulavski *et al.* (1986) *J. Immunol. Methods* 89:123-130; Prooijen-Knegt (1982) *Exp. Cell Res.* 141:397-407; Rudkin (1976) *Nature* 265:472-473, Stollar (1970) *PNAS* 65:993-1000; Ballard (1982) *Mol. Immunol.* 19:793-799; Pisetsky and Caster (1982) *Mol. Immunol.* 19:645-650; Viscidi *et al.* (1988) *J. Clin. Microbial.* 41:199-209, and Kiney *et al.* (1989) *J. Clin. Microbiol.* 27:6-12 describe antibodies to RNA duplexes, including homo and heteroduplexes. Kits comprising antibodies specific for DNA:RNA hybrids are available, *e.g.*, from Digene Diagnostics, Inc. (Beltsville, MD).

In addition to available antibodies, one of skill in the art can easily make antibodies specific for nucleic acid duplexes using existing techniques, or modify those antibodies which are commercially or publicly available. In addition to the art referenced above, general methods of producing polyclonal and monoclonal antibodies are known to those of skill in the art. *See, e.g.*, Paul (ed)

(1993) *Fundamental Immunology, Third Edition* Raven Press, Ltd., New York Coligan (1991) *Current Protocols in Immunology* Wiley/Greene, NY; Harlow and Lane (1989) *Antibodies: A Laboratory Manual* Cold Spring Harbor Press, NY; Stites *et al.* (eds.) *Basic and Clinical Immunology* (4th ed.) Lange Medical Publications, Los Altos, CA, and references cited therein; Goding (1986) *Monoclonal Antibodies: Principles and Practice* (2d ed.) Academic Press, New York, NY; and Kohler and Milstein (1975) *Nature* 256: 495-497. Other suitable techniques for antibody preparation include selection of libraries of recombinant antibodies in phage or similar vectors. *See*, Huse *et al.* (1989) *Science* 246: 1275-1281; and Ward *et al.* (1989) *Nature* 341: 544-546. Specific monoclonal and polyclonal antibodies and antisera will usually bind with a K_D of at least about 0.1 μM, preferably at least about 0.01 μM or better, and most typically and preferably, 0.001 μM or better.

The nucleic acids used in this invention can be either positive or negative probes. Positive probes bind to their targets and the presence of duplex formation is evidence of the presence of the target. Negative probes fail to bind to the suspect target and the absence of duplex formation is evidence of the presence of the target. For example, the use of a wild type specific nucleic acid probe or PCR primers may serve as a negative probe in an assay sample where only the nucleotide sequence of interest is present.

The sensitivity of the hybridization assays may be enhanced through use of a nucleic acid amplification system which multiplies the target nucleic acid being detected. Examples of such systems include the polymerase chain reaction (PCR) system and the ligase chain reaction (LCR) system. Other methods recently described in the art are the nucleic acid sequence based amplification (NASBA®, Cangene, Mississauga, Ontario) and Q Beta Replicase systems. These systems can be used to directly identify mutants where the PCR or LCR primers are designed to be extended or ligated only when a selected sequence is present. Alternatively, the selected sequences can be generally amplified using, for example, nonspecific PCR primers and the amplified target region later probed for a specific sequence indicative of a mutation.

A preferred embodiment is the use of allelic specific amplifications. In the case of PCR, the amplification primers are designed to bind to a portion of,

for example, a gene encoding a cancer-associated protein, but the terminal base at the 3' end is used to discriminate between the mutant and wild-type forms of the cancer-associated protein gene. If the terminal base matches the point mutation or the wild-type, polymerase dependent three prime extension can proceed and an amplification product is detected. This method for detecting point mutations or polymorphisms is described in detail by Sommer, S.S., *et al.*, in *Mayo Clin. Proc.* 64:1361-1372 (1989), incorporated herein by reference. By using appropriate controls, one can develop a kit having both positive and negative amplification products. The products can be detected using specific probes or by simply detecting their presence or absence. A variation of the PCR method uses LCR where the point of discrimination, i.e, either the point mutation or the wild-type bases fall between the LCR oligonucleotides. The ligation of the oligonucleotides becomes the means for discriminating between the mutant and wild-type forms of the cancer-associated protein gene.

An alternative means for determining the level of expression of the nucleic acids of the present invention is *in situ* hybridization. *In situ* hybridization assays are well known and are generally described in Angerer, *et al., Methods Enzymol.*, 152:649-660 (1987). In an *in situ* hybridization assay cells, preferentially bovine lymphocytes, are fixed to a solid support, typically a glass slide. If DNA is to be probed, the cells are denatured with heat or alkali. The cells are then contacted with a hybridization solution at a moderate temperature to permit annealing of specific probes that are labeled. The probes are preferably labeled with radioisotopes or fluorescent reporters.

II. Immunological Detection of a Cancer-Associated Protein

In addition to the detection of the subject protein gene expression using nucleic acid hybridization technology, one can also use immunoassays to detect the protein itself. Immunoassays can be used to qualitatively or quantitatively analyze the proteins of interest. A general overview of the applicable technology can be found in Harlow and Lane, *Antibodies: A Laboratory Manual,* Cold Spring Harbor Pubs., N.Y. (1988), incorporated herein by reference. Although the following discussion is directed to methods for detecting target proteins associated

with cancer, similar methods can be used to detect methods associated with cell proliferation, cell transformation, neoplasia, metastasis and/or target proteins associated with different types of cancer (*e.g.*, liver, kidney, breast, prostate, colon, *etc.*).

A. Antibodies to Target Proteins

Methods of producing polyclonal and monoclonal antibodies that react specifically with a protein of interest are known to those of skill in the art. *See, e.g.,* Coligan (1991), Current Protocols in Immunology, Wiley/Greene, NY; and Harlow and Lane; Stites *et al.* (eds.) Basic and Clinical Immunology (4th ed.) Lange Medical Publications, Los Altos, CA, and references cited therein; Goding (1986), Monoclonal antibodies: Principles and Practice (2d ed.) Academic Press, New York, NY; and Kohler and Milstein (1975) *Nature*, 256:495-497. Such techniques include antibody preparation by selection of antibodies from libraries of recombinant antibodies in phage or similar vectors. *See*, Huse *et al.* (1989) *Science*, 246:1275-1281; and Ward *et al.* (1989) *Nature*, 341:544-546. For example, in order to produce antisera for use in an immunoassay, the proteins of interest or an antigenic fragment thereof, is isolated as described herein. For example, recombinant protein is produced in a transformed cell line. An inbred strain of mice or rabbits is immunized with the protein using a standard adjuvant, such as Freund's adjuvant, and a standard immunization protocol. Alternatively, a synthetic peptide derived from the sequences disclosed herein and conjugated to a carrier protein can be used as an immunogen.

Polyclonal sera are collected and titered against the immunogen protein in an immunoassay, for example, a solid phase immunoassay with the immunogen immobilized on a solid support. Polyclonal antisera with a titer of 10^4 or greater are selected and tested for their cross reactivity against non-cancer-associated proteins or even other homologous proteins from other organisms, using a competitive binding immunoassay. Specific monoclonal and polyclonal antibodies and antisera will usually bind with a K_D of at least about 0.1 mM, more usually at least about 1 μM, preferably at least about 0.1 μM or better, and most preferably, 0.01 μM or better.

A number of proteins of the invention comprising immunogens may be used to produce antibodies specifically or selectively reactive with the proteins of interest. Recombinant protein is the preferred immunogen for the production of monoclonal or polyclonal antibodies. Naturally occurring protein may also be used either in pure or impure form. Synthetic peptides made using the protein sequences described herein may also be used as an immunogen for the production of antibodies to the protein. Recombinant protein can be expressed in eukaryotic or prokaryotic cells and purified as generally described *infra*. The product is then injected into an animal capable of producing antibodies. Either monoclonal or polyclonal antibodies may be generated, for subsequent use in immunoassays to measure the protein.

Methods of production of polyclonal antibodies are known to those of skill in the art. In brief, an immunogen, preferably a purified protein, is mixed with an adjuvant and animals are immunized. The animal's immune response to the immunogen preparation is monitored by taking test bleeds and determining the titer of reactivity to a cancer-associated protein. When appropriately high titers of antibody to the immunogen are obtained, blood is collected from the animal and antisera are prepared. Further fractionation of the antisera to enrich for antibodies reactive to the protein can be done if desired (*see*, Harlow and Lane, *supra*).

Monoclonal antibodies may be obtained by various techniques familiar to those skilled in the art. Briefly, spleen cells from an animal immunized with a desired antigen are immortalized, commonly by fusion with a myeloma cell (*See*, Kohler and Milstein, *Eur. J. Immunol.* 6:511-519 (1976), incorporated herein by reference). Alternative methods of immortalization include transformation with Epstein Barr Virus, oncogenes, or retroviruses, or other methods well known in the art. Colonies arising from single immortalized cells are screened for production of antibodies of the desired specificity and affinity for the antigen, and yield of the monoclonal antibodies produced by such cells may be enhanced by various techniques, including injection into the peritoneal cavity of a vertebrate host. Alternatively, one may isolate DNA sequences which encode a monoclonal antibody or a binding fragment thereof by screening a DNA library from human B cells according to the general protocol outlined by Huse, *et al.* (1989) *Science* 246:1275-1281.

Once target protein specific antibodies are available, the protein can be measured by a variety of immunoassay methods with qualitative and quantitative results available to the clinician. For a review of immunological and immunoassay procedures in general *see, Basic and Clinical Immunology* 7th Edition (D. Stites and A. Terr ed.) 1991. Moreover, the immunoassays of the present invention can be performed in any of several configurations, which are reviewed extensively in *Enzyme Immunoassay*, E.T. Maggio, ed., CRC Press, Boca Raton, Florida (1980); "Practice and Theory of Enzyme Immunoassays," Tijssen; and, Harlow and Lane, each of which is incorporated herein by reference.

Immunoassays to measure target proteins in a human sample may use a polyclonal antiserum which was raised to the protein partially encoded by a sequence described herein or a fragment thereof. This antiserum is selected to have low crossreactivity against non-cancer-associated proteins and any such crossreactivity is removed by immunoabsorption prior to use in the immunoassay.

In order to produce antisera for use in an immunoassay, cancer-associated protein or a fragment thereof, for example, is isolated as described herein. For example, recombinant protein is produced in a transformed cell line. An inbred strain of mice, such as Balb/c, is immunized with the protein or a peptide using a standard adjuvant, such as Freund's adjuvant, and a standard mouse immunization protocol. Alternatively, a synthetic peptide derived from the sequences disclosed herein and conjugated to a carrier protein can be used an immunogen. Polyclonal sera are collected and titered against the immunogen protein in an immunoassay, for example, a solid phase immunoassay with the immunogen immobilized on a solid support. Polyclonal antisera with a titer of 10^4 or greater are selected and tested for their cross reactivity against non-cancer-associated proteins, using a competitive binding immunoassay such as the one described in Harlow and Lane, *supra*, at pages 570-573 and below.

B. Immunological Binding Assays

In a preferred embodiment, a protein of interest is detected and/or quantified using any of a number of well recognized immunological binding assays (*see, e.g.,* U.S. Patents 4,366,241; 4,376,110; 4,517,288; and 4,837,168). For a review of the general immunoassays, see also *Methods in Cell Biology Volume 37:*

Antibodies in Cell Biology, Asai, ed. Academic Press, Inc. New York (1993); *Basic and Clinical Immunology* 7th Edition, Stites & Terr, eds. (1991). Immunological binding assays (or immunoassays) typically utilize a "capture agent" to specifically bind to and often immobilize the analyte (in this case the cancer-associated protein or antigenic subsequence thereof). The capture agent is a moiety that specifically binds to the analyte. In a preferred embodiment, the capture agent is an antibody that specifically binds, for example, the cancer-associated protein. The antibody (*e.g.,* anti-cancer-associated protein antibody) may be produced by any of a number of means well known to those of skill in the art and as described above.

Immunoassays also often utilize a labeling agent to specifically bind to and label the binding complex formed by the capture agent and the analyte. The labeling agent may itself be one of the moieties comprising the antibody/analyte complex. Thus, the labeling agent may be a labeled cancer-associated protein polypeptide or a labeled anti-cancer-associated protein antibody. Alternatively, the labeling agent may be a third moiety, such as another antibody, that specifically binds to the antibody/protein complex.

In a preferred embodiment, the labeling agent is a second antibody bearing a label. Alternatively, the second antibody may lack a label, but it may, in turn, be bound by a labeled third antibody specific to antibodies of the species from which the second antibody is derived. The second antibody can be modified with a detectable moiety, such as biotin, to which a third labeled molecule can specifically bind, such as enzyme-labeled streptavidin.

Other proteins capable of specifically binding immunoglobulin constant regions, such as protein A or protein G, can also be used as the label agent. These proteins are normal constituents of the cell walls of streptococcal bacteria. They exhibit a strong non-immunogenic reactivity with immunoglobulin constant regions from a variety of species (*see,* generally, Kronval, *et al.* (1973) *J. Immunol.*, 111: 1401-1406, and Akerstrom, *et al.* (1985) *J. Immunol.*, 135: 2589-2542).

Throughout the assays, incubation and/or washing steps may be required after each combination of reagents. Incubation steps can vary from about 5 seconds to several hours, preferably from about 5 minutes to about 24 hours. However, the incubation time will depend upon the assay format, analyte, volume of solution, concentrations, and the like. Usually, the assays will be carried out at

ambient temperature, although they can be conducted over a range of temperatures, such as 10°C to 40°C.

1. Non-Competitive Assay Formats

Immunoassays for detecting proteins of interest from tissue samples may be either competitive or noncompetitive. Noncompetitive immunoassays are assays in which the amount of captured analyte (in this case the protein) is directly measured. In one preferred "sandwich" assay, for example, the capture agent (*e.g.*, anti-cancer-associated protein antibodies) can be bound directly to a solid substrate where it is immobilized. These immobilized antibodies then capture the cancer-associated protein present in the test sample. The cancer-associated protein thus immobilized is then bound by a labeling agent, such as a second cancer-associated protein antibody bearing a label. Alternatively, the second antibody may lack a label, but it may, in turn, be bound by a labeled third antibody specific to antibodies of the species from which the second antibody is derived. The second can be modified with a detectable moiety, such as biotin, to which a third labeled molecule can specifically bind, such as enzyme-labeled streptavidin.

2. Competitive Assay Formats

In competitive assays, the amount of target protein (analyte) present in the sample is measured indirectly by measuring the amount of an added (exogenous) analyte (*i.e.*, the cancer-associated protein of interest) displaced (or competed away) from a capture agent (anti-cancer-associated protein antibody) by the analyte present in the sample. In one competitive assay, a known amount of, in this case, the protein of interest is added to the sample and the sample is then contacted with a capture agent, in this case an antibody that specifically binds to the cancer-associated protein. The amount of cancer-associated protein bound to the antibody is inversely proportional to the concentration of cancer-associated protein present in the sample. In a particularly preferred embodiment, the antibody is immobilized on a solid substrate. The amount of the cancer-associated protein bound to the antibody may be determined either by measuring the amount of subject protein present in a cancer-associated protein/antibody complex or, alternatively, by measuring the amount of remaining uncomplexed protein. The amount of cancer-

associated protein may be detected by providing a labeled cancer-associated protein molecule.

A hapten inhibition assay is another preferred competitive assay. In this assay, a known analyte, in this case the target protein, is immobilized on a solid substrate. A known amount of anti-cancer-associated protein antibody is added to the sample, and the sample is then contacted with the immobilized target. In this case, the amount of anti-cancer-associated protein antibody bound to the immobilized cancer-associated protein is inversely proportional to the amount of cancer-associated protein present in the sample. Again, the amount of immobilized antibody may be detected by detecting either the immobilized fraction of antibody or the fraction of the antibody that remains in solution. Detection may be direct where the antibody is labeled or indirect by the subsequent addition of a labeled moiety that specifically binds to the antibody as described above.

Immunoassays in the competitive binding format can be used for crossreactivity determinations. For example, the protein encoded by the sequences described herein can be immobilized to a solid support. Proteins are added to the assay which compete with the binding of the antisera to the immobilized antigen. The ability of the above proteins to compete with the binding of the antisera to the immobilized protein is compared to the protein encoded by any of the sequences described herein. The percent crossreactivity for the above proteins is calculated, using standard calculations. Those antisera with less than 10% crossreactivity with each of the proteins listed above are selected and pooled. The cross-reacting antibodies are optionally removed from the pooled antisera by immunoabsorption with the considered proteins, *e.g.*, distantly related homologues.

The immunoabsorbed and pooled antisera are then used in a competitive binding immunoassay as described above to compare a second protein, thought to be perhaps the protein of this invention, to the immunogen protein. In order to make this comparison, the two proteins are each assayed at a wide range of concentrations and the amount of each protein required to inhibit 50% of the binding of the antisera to the immobilized protein is determined. If the amount of the second protein required is less than 10 times the amount of the protein partially encoded by a sequence herein that is required, then the second protein is said to specifically bind to an antibody generated to an immunogen consisting of the target protein.

3. Other Assay Formats

In a particularly preferred embodiment, Western blot (immunoblot) analysis is used to detect and quantify the presence of cancer-associated protein in the sample. The technique generally comprises separating sample proteins by gel electrophoresis on the basis of molecular weight, transferring the separated proteins to a suitable solid support (such as a nitrocellulose filter, a nylon filter, or derivatized nylon filter) and incubating the sample with the antibodies that specifically bind the protein of interest. For example, the anti-cancer-associated protein antibodies specifically bind to the cancer-associated protein on the solid support. These antibodies may be directly labeled or alternatively may be subsequently detected using labeled antibodies (*e.g.*, labeled sheep anti-mouse antibodies) that specifically bind to the antibodies against the protein of interest.

Other assay formats include liposome immunoassays (LIA), which use liposomes designed to bind specific molecules (*e.g.*, antibodies) and release encapsulated reagents or markers. The released chemicals are then detected according to standard techniques (*see*, Monroe *et al.* (1986) *Amer. Clin. Prod. Rev.* 5:34-41).

4. Reduction of Non-Specific Binding

One of skill in the art will appreciate that it is often desirable to use non-specific binding in immunoassays. Particularly, where the assay involves an antigen or antibody immobilized on a solid substrate it is desirable to minimize the amount of non-specific binding to the substrate. Means of using such non-specific binding are well known to those of skill in the art. Typically, this involves coating the substrate with a proteinaceous composition. In particular, protein compositions, such as bovine serum albumin (BSA), nonfat powdered milk and gelatin, are widely used with powdered milk being most preferred.

5. Labels

The particular label or detectable group used in the assay is not a critical aspect of the invention, so long as it does not significantly interfere with the specific binding of the antibody used in the assay. The detectable group can be any material having a detectable physical or chemical property. Such detectable labels

have been well-developed in the field of immunoassays and, in general, most labels useful in such methods can be applied to the present invention. Thus, a label is any composition detectable by spectroscopic, photochemical, biochemical, immunochemical, electrical, optical or chemical means. Useful labels in the present invention include magnetic beads (*e.g.*, Dynabeads^{TM}), fluorescent dyes (*e.g.*, fluorescein isothiocyanate, Texas red, rhodamine, and the like), radiolabels (*e.g.*, ^{3}H, ^{125}I, ^{35}S, ^{14}C, or ^{32}P), enzymes (*e.g.*, horse radish peroxidase, alkaline phosphatase and others commonly used in an ELISA), and colorimetric labels such as colloidal gold or colored glass or plastic (*e.g.*, polystyrene, polypropylene, latex, *etc.*) beads.

The label may be coupled directly or indirectly to the desired component of the assay according to methods well known in the art. As indicated above, a wide variety of labels may be used, with the choice of label depending on sensitivity required, ease of conjugation with the compound, stability requirements, available instrumentation, and disposal provisions.

Non-radioactive labels are often attached by indirect means. Generally, a ligand molecule (*e.g.*, biotin) is covalently bound to the molecule. The ligand then binds to an anti-ligand (*e.g.*, streptavidin) molecule which is either inherently detectable or covalently bound to a signal system, such as a detectable enzyme, a fluorescent compound, or a chemiluminescent compound. A number of ligands and anti-ligands can be used. Thyroxine, and cortisol can be used in conjunction with the labeled, naturally occurring anti-ligands. Alternatively, any haptenic or antigenic compound can be used in combination with an antibody.

The molecules can also be conjugated directly to signal generating compounds, *e.g.*, by conjugation with an enzyme or fluorophore. Enzymes of interest as labels will primarily be hydrolases, particularly phosphatases, esterases and glycosidases, or oxidotases, particularly peroxidases. Fluorescent compounds include fluorescein and its derivatives, rhodamine and its derivatives, dansyl, umbelliferone, *etc.* Chemiluminescent compounds include luciferin, and 2,3-dihydrophthalazinediones, *e.g.*, luminol. For a review of various labeling or signal producing systems which may be used, see, U.S. Patent No. 4,391,904).

Means of detecting labels are well known to those of skill in the art. Thus, for example, where the label is a radioactive label, means for detection include a scintillation counter or photographic film as in autoradiography. Where the label

is a fluorescent label, it may be detected by exciting the fluorochrome with the appropriate wavelength of light and detecting the resulting fluorescence. The fluorescence may be detected visually, by means of photographic film, by the use of electronic detectors such as charge coupled devices (CCDs) or photomultipliers and the like. Similarly, enzymatic labels may be detected by providing the appropriate substrates for the enzyme and detecting the resulting reaction product. Finally simple colorimetric labels may be detected directly by observing the color associated with the label. Thus, in various dipstick assays, conjugated gold often appears pink, while various conjugated beads appear the color of the bead.

Some assay formats do not require the use of labeled components. For instance, agglutination assays can be used to detect the presence of the target antibodies. In this case, antigen-coated particles are agglutinated by samples comprising the target antibodies. In this format, none of the components need be labeled and the presence of the target antibody is detected by simple visual inspection.

III. Screening for Modulators of the Cancer Development Process

The invention also provides methods for identifying compounds that modulate cancer development, *e.g.*, loss of proliferation control, cell transformation, development of malignancy and/or metastasis. For example, the methods can identify compounds that increase or decrease the expression level of genes associated with cancer (*e.g.*, cell proliferation, neoplasia, cell transformation, malignancy, metastatic ability, *etc.*) and cancer-related conditions. Although the following discussion is directed to methods for screening for modulators of cancer development, similar methods can be used to screen for modulators of cell proliferation, neoplasia, cell transformation or for modulators of different types of cancer (*e.g.*, liver, kidney, breast, prostate, colon, *etc.*).

For instance, compounds that are identified as modulators of cancer development using the methods of the invention find use both *in vitro* and *in vivo*. For example, one can treat cell cultures with the modulators in experiments designed to determine the mechanisms by which cancer development is regulated. Compounds that increase tumor development and/or induce neoplasia are useful for

extending the life of cell cultures that are used for production of biological products such as recombinant proteins. *In vivo* uses of compounds that delay cancer development include, for example, delaying the loss of cell proliferation control and/or cell transformation and treating conditions associated with neoplasia, malignant tumor formation and/or metastasis.

The methods typically involve culturing a cell in the presence of a potential modulator to form a first cell culture. RNA (or cDNA) from the first cell culture is contacted with a probe which comprises a polynucleotide sequence associated with cancer. The amount of the probe which hybridizes to the RNA (or cDNA) from the first cell culture is determined. Typically, one determines whether the amount of probe which hybridizes to the RNA (or cDNA) is increased or decreased relative to the amount of the probe which hybridizes to RNA (or cDNA) from a second cell culture grown in the absence of the modulator.

It may be further determined whether the modulator-induced increase or decrease in RNA (or cDNA) levels of the target sequence is correlated with any cancer-associated change in cellular phenotype. For example, a liver, kidney, breast, colon or prostate cell population that is treated with a modulator which induces decreased expression of a gene that is normally upregulated with cancer, or a liver, kidney, breast, colon or prostate cell that is treated with a modulator which induces increased expression of a gene that is normally downregulated with cancer, may be further tested for loss of control of cell proliferation, which is reflective, for example, of a neoplastic phenotype, for cell transformation or for metastatic ability.

Essentially any chemical compound can be used as a potential modulator in the assays of the invention, although most often compounds can be dissolved in aqueous or organic (for example, DMSO-based) solutions are used. The assays are designed to screen large chemical libraries by automating the assay steps and providing compounds from any convenient source to assays, which are typically run in parallel (*e.g.*, in microtiter formats on microtiter plates in robotic assays). It will be appreciated that there are many suppliers of chemical compounds, including Sigma (St. Louis, MO), Aldrich (St. Louis, MO), Sigma-Aldrich (St. Louis, MO), Fluka Chemika-Biochemica Analytika (Buchs Switzerland) and the like.

In one preferred embodiment, high throughput screening methods involve providing a combinatorial library containing a large number of potential therapeutic compounds (potential modulator compounds). Such "combinatorial chemical libraries" are then screened in one or more assays, as described herein, to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional "lead compounds" or can themselves be used as potential or actual therapeutics.

A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis, by combining a number of chemical "building blocks" such as reagents. For example, a linear combinatorial chemical library such as a polypeptide library is formed by combining a set of chemical building blocks (amino acids) in every possible way for a given compound length (*i.e.*, the number of amino acids in a polypeptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks.

Preparation and screening of combinatorial chemical libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to, peptide libraries (*see, e.g.,* U.S. Patent 5,010,175, Furka, *Int. J. Pept. Prot. Res.* 37:487-493 (1991) and Houghton *et al.*, *Nature* 354:84-88 (1991)). Other chemistries for generating chemical diversity libraries can also be used. Such chemistries include, but are not limited to: peptoids (PCT Publication No. WO 91/19735), encoded peptides (PCT Publication WO 93/20242), random bio-oligomers (PCT Publication No. WO 92/00091), benzodiazepines (U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs *et al., Proc. Nat. Acad. Sci. USA* 90:6909-6913 (1993)), vinylogous polypeptides (Hagihara *et al.*, *J. Amer. Chem. Soc.* 114:6568 (1992)), nonpeptidal peptidomimetics with β-D-glucose scaffolding (Hirschmann *et al.*, *J. Amer. Chem. Soc.* 114:9217-9218 (1992)), analogous organic syntheses of small compound libraries (Chen *et al.*, *J. Amer. Chem. Soc.* 116:2661 (1994)), oligocarbamates (Cho *et al.*, *Science* 261:1303 (1993)), and/or peptidyl phosphonates (Campbell *et al., J. Org. Chem.* 59:658 (1994)), nucleic acid libraries (*see,* Ausubel, Berger and Sambrook, *all supra*), peptide nucleic acid libraries (*see, e.g.,* U.S. Patent

5,539,083), antibody libraries (*see, e.g.,* Vaughn *et al.*, *Nature Biotechnology*, 14(3):309-314 (1996) and PCT/US96/10287), carbohydrate libraries (*see, e.g.,* Liang *et al.*, *Science*, 274:1520-1522 (1996) and U.S. Patent 5,593,853), small organic molecule libraries (*see, e.g.,* benzodiazepines, Baum C&EN, Jan 18, page 33 (1993); isoprenoids, U.S. Patent 5,569,588; thiazolidinones and metathiazanones, U.S. Patent 5,549,974; pyrrolidines, U.S. Patents 5,525,735 and 5,519,134; morpholino compounds, U.S. Patent 5,506,337; benzodiazepines, 5,288,514, and the like.

Devices for the preparation of combinatorial libraries are commercially available (*see, e.g.,* 357 MPS, 390 MPS, Advanced Chem Tech, Louisville KY, Symphony, Rainin, Woburn, MA, 433A Applied Biosystems, Foster City, CA, 9050 Plus, Millipore, Bedford, MA). In addition, numerous combinatorial libraries are themselves commercially available (*see, e.g.,* ComGenex, Princeton, N.J., Asinex, Moscow, Ru, Tripos, Inc., St. Louis, MO, ChemStar, Ltd, Moscow, RU, 3D Pharmaceuticals, Exton, PA, Martek Biosciences, Columbia, MD, *etc.*).

As noted, the invention provides *in vitro* assays for identifying, in a high throughput format, compounds that can modulate cancer development. Control reactions that measure the transformation or malignancy level of the cell in a reaction that does not include a potential modulator are optional, as the assays are highly uniform. Such optional control reactions are appropriate and increase the reliability of the assay. Accordingly, in a preferred embodiment, the methods of the invention include such a control reaction. For each of the assay formats described, "no modulator" control reactions which do not include a modulator provide a background level of binding activity.

In some assays it will be desirable to have positive controls to ensure that the components of the assays are working properly. At least two types of positive controls are appropriate. First, a known activator of cancer development can be incubated with one sample of the assay, and the resulting increase in signal resulting from an increased expression level of a gene associated with cancer determined according to the methods herein. Second, a known inhibitor of cancer development can be added, and the resulting decrease in signal for the expression of a gene associated with cancer similarly detected. It will be appreciated that modulators can also be combined with activators or inhibitors to find modulators

which inhibit the increase or decrease that is otherwise caused by the presence of the known modulator of cancer development.

In the high throughput assays of the invention, it is possible to screen up to several thousand different modulators in a single day. In particular, each well of a microtiter plate can be used to run a separate assay against a selected potential modulator, or, if concentration or incubation time effects are to be observed, every 5-10 wells can test a single modulator. Thus, a single standard microtiter plate can assay about 100 (96) modulators. If 1536 well plates are used, then a single plate can easily assay from about 100- about 1500 different compounds. It is possible to assay many different plates per day; assay screens for up to about 6,000-20,000, and even up to about 100,000 different compounds is possible using the integrated systems of the invention.

IV. Compositions, Kits and Integrated Systems

The invention provides compositions, kits and integrated systems for practicing the assays described herein. Although the following discussion is directed to kits for carrying out assays using nucleic acids (or proteins, antibodies, *etc.*) associated with cancer, similar kits can be assembled for carrying out assays using nucleic acids (or proteins, antibodies, *etc.*) associated with cell proliferation, neoplasia, cell transformation, malignancy, metastasis and/or nucleic acids associated with different types of cancer (*e.g.*, liver, kidney, breast, prostate, colon, *etc.*). For instance, an assay composition having a nucleic acid associated with, for example, malignancy of a cell and a labeling reagent is provided by the present invention. In preferred embodiments, a plurality of, for example, cancer-associated nucleic acids are provided in the assay compositions. The invention also provides assay compositions for use in solid phase assays; such compositions can include, for example, one or more cancer-associated nucleic acids immobilized on a solid support, and a labeling reagent. In each case, the assay compositions can also include additional reagents that are desirable for hybridization. Modulators of expression of, for example, cancer-associated nucleic acids can also be included in the assay compositions.

The invention also provides kits for carrying out the assays of the invention. The kits typically include a probe which comprises a polynucleotide sequence associated with cancer; and a label for detecting the presence of the probe. Preferably, the kits will include a plurality of polynucleotide sequences associated with cancer. Kits can include any of the compositions noted above, and optionally further include additional components such as instructions to practice a high-throughput method of assaying for an effect on cell proliferation and transformation and expression of cancer-associated genes, one or more containers or compartments (*e.g.*, to hold the probe, labels, or the like), a control modulator of cancer development, a robotic armature for mixing kit components or the like.

The invention also provides integrated systems for high-throughput screening of potential modulators for an effect on cancer development. The systems typically include a robotic armature which transfers fluid from a source to a destination, a controller which controls the robotic armature, a label detector, a data storage unit which records label detection, and an assay component such as a microtiter dish comprising a well having a reaction mixture or a substrate comprising a fixed nucleic acid or immobilization moiety.

A number of robotic fluid transfer systems are available, or can easily be made from existing components. For example, a Zymate XP (Zymark Corporation; Hopkinton, MA) automated robot using a Microlab 2200 (Hamilton; Reno, NV) pipetting station can be used to transfer parallel samples to 96 well microtiter plates to set up several parallel simultaneous STAT binding assays.

Optical images viewed (and, optionally, recorded) by a camera or other recording device (*e.g.*, a photodiode and data storage device) are optionally further processed in any of the embodiments herein, *e.g.*, by digitizing the image and storing and analyzing the image on a computer. A variety of commercially available peripheral equipment and software is available for digitizing, storing and analyzing a digitized video or digitized optical image, *e.g.*, using PC (Intel x86 or Pentium chip-compatible DOS®, OS2® WINDOWS®, WINDOWS NT® or WINDOWS95® based computers), MACINTOSH®, or UNIX® based (*e.g.*, SUN® work station) computers.

One conventional system carries light from the specimen field to a cooled charge-coupled device (CCD) camera, in common use in the art. A CCD camera includes an array of picture elements (pixels). The light from the specimen

is imaged on the CCD. Particular pixels corresponding to regions of the specimen (*e.g.*, individual hybridization sites on an array of biological polymers) are sampled to obtain light intensity readings for each position. Multiple pixels are processed in parallel to increase speed. The apparatus and methods of the invention are easily used for viewing any sample, *e.g.*, by fluorescent or dark field microscopic techniques.

V. Gene Therapy Applications

A variety of human diseases can be treated by therapeutic approaches that involve stably introducing a gene into a human cell such that the gene is transcribed and the gene product is produced in the cell. Diseases amenable to treatment by this approach include inherited diseases, including those in which the defect is in a single gene. Gene therapy is also useful for treatment of acquired diseases and other conditions. For discussions on the application of gene therapy towards the treatment of genetic as well as acquired diseases, *see*, Miller, A.D. (1992) *Nature* 357:455-460, and Mulligan, R.C. (1993) *Science* 260:926-932, both of which are incorporated herein by reference.

A. Vectors for Gene Delivery

For delivery to a cell or organism, the nucleic acids of the invention can be incorporated into a vector. Examples of vectors used for such purposes include expression plasmids capable of directing the expression of the nucleic acids in the target cell. In other instances, the vector is a viral vector system wherein the nucleic acids are incorporated into a viral genome that is capable of transfecting the target cell. In a preferred embodiment, the nucleic acids can be operably linked to expression and control sequences that can direct expression of the gene in the desired target host cells. Thus, one can achieve expression of the nucleic acid under appropriate conditions in the target cell.

B. Gene Delivery Systems

Viral vector systems useful in the expression of the nucleic acids include, for example, naturally occurring or recombinant viral vector systems.

Depending upon the particular application, suitable viral vectors include replication competent, replication deficient, and conditionally replicating viral vectors. For example, viral vectors can be derived from the genome of human or bovine adenoviruses, vaccinia virus, herpes virus, adeno-associated virus, minute virus of mice (MVM), HIV, sindbis virus, and retroviruses (including but not limited to Rous sarcoma virus), and MoMLV. Typically, genes of interest are inserted into such vectors to allow packaging of the gene construct, typically with accompanying viral DNA, followed by infection of a sensitive host cell and expression of the gene of interest.

As used herein, "gene delivery system" refers to any means for the delivery of a nucleic acid of the invention to a target cell. In some embodiments of the invention, nucleic acids are conjugated to a cell receptor ligand for facilitated uptake (*e.g.*, invagination of coated pits and internalization of the endosome) through an appropriate linking moiety, such as a DNA linking moiety (Wu *et al.*, *J. Biol. Chem.* 263:14621-14624 (1988); WO 92/06180). For example, nucleic acids can be linked through a polylysine moiety to asialo-oromucocid, which is a ligand for the asialoglycoprotein receptor of hepatocytes.

Similarly, viral envelopes used for packaging gene constructs that include the nucleic acids of the invention can be modified by the addition of receptor ligands or antibodies specific for a receptor to permit receptor-mediated endocytosis into specific cells (*see, e.g.*, WO 93/20221, WO 93/14188, WO 94/06923). In some embodiments of the invention, the DNA constructs of the invention are linked to viral proteins, such as adenovirus particles, to facilitate endocytosis (Curiel *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 88: 8850-8854 (1991)). In other embodiments, molecular conjugates of the instant invention can include microtubule inhibitors (WO/9406922); synthetic peptides mimicking influenza virus hemagglutinin (Plank *et al.*, *J. Biol. Chem.* 269:12918-12924 (1994)); and nuclear localization signals such as SV40 T antigen (WO93/19768).

Retroviral vectors are also useful for introducing the nucleic acids of the invention into target cells or organisms. Retroviral vectors are produced by genetically manipulating retroviruses. The viral genome of retroviruses is RNA. Upon infection, this genomic RNA is reverse transcribed into a DNA copy which is integrated into the chromosomal DNA of transduced cells with a high degree of

stability and efficiency. The integrated DNA copy is referred to as a provirus and is inherited by daughter cells as is any other gene. The wild type retroviral genome and the proviral DNA have three genes: the *gag*, the *pol* and the *env* genes, which are flanked by two long terminal repeat (LTR) sequences. The *gag* gene encodes the internal structural (nucleocapsid) proteins; the *pol* gene encodes the RNA directed DNA polymerase (reverse transcriptase); and the *env* gene encodes viral envelope glycoproteins. The 5' and 3' LTRs serve to promote transcription and polyadenylation of virion RNAs. Adjacent to the 5' LTR are sequences necessary for reverse transcription of the genome (the tRNA primer binding site) and for efficient encapsulation of viral RNA into particles (the Psi site). *See*, Mulligan, R.C., In: *Experimental Manipulation of Gene Expression*, M. Inouye (ed), 155-173 (1983); Mann, R., *et al.*, *Cell*, 33:153-159 (1983); Cone, R.D. and R.C. Mulligan, *Proceedings of the National Academy of Sciences, U.S.A.*, 81:6349-6353 (1984).

The design of retroviral vectors is well known to those of ordinary skill in the art. *See, e.g.*, Singer, M. and Berg, P., *supra*. In brief, if the sequences necessary for encapsidation (or packaging of retroviral RNA into infectious virions) are missing from the viral genome, the result is a *cis* acting defect which prevents encapsidation of genomic RNA. However, the resulting mutant is still capable of directing the synthesis of all virion proteins. Retroviral genomes from which these sequences have been deleted, as well as cell lines containing the mutant genome stably integrated into the chromosome are well known in the art and are used to construct retroviral vectors. Preparation of retroviral vectors and their uses are described in many publications including European Patent Application EPA 0 178 220, U.S. Patent 4,405,712, Gilboa, *Biotechniques* 4:504-512 (1986), Mann, *et al.*, *Cell* 33:153-159 (1983), Cone and Mulligan, *Proc. Natl. Acad. Sci. USA* 81:6349-6353 (1984), Eglitis, M.A, *et al.* (1988) *Biotechniques* 6:608-614, Miller, A.D. *et al.* (1989) *Biotechniques* 7:981-990, Miller, A.D.(1992) *Nature*, *supra*, Mulligan, R.C. (1993), *supra*, and Gould, B. *et al.*, and International Publication No. WO 92/07943 entitled "Retroviral Vectors Useful in Gene Therapy". The teachings of these patents and publications are incorporated herein by reference.

The retroviral vector particles are prepared by recombinantly inserting the desired nucleotide sequence into a retrovirus vector and packaging the vector with retroviral capsid proteins by use of a packaging cell line. The resultant

retroviral vector particle is incapable of replication in the host cell but is capable of integrating into the host cell genome as a proviral sequence containing the desired nucleotide sequence. As a result, the patient is capable of producing, for example, the cancer-associated protein and thus restore the cells to a normal, non-cancerous phenotype.

Packaging cell lines that are used to prepare the retroviral vector particles are typically recombinant mammalian tissue culture cell lines that produce the necessary viral structural proteins required for packaging, but which are incapable of producing infectious virions. The defective retroviral vectors that are used, on the other hand, lack these structural genes but encode the remaining proteins necessary for packaging. To prepare a packaging cell line, one can construct an infectious clone of a desired retrovirus in which the packaging site has been deleted. Cells comprising this construct will express all structural viral proteins, but the introduced DNA will be incapable of being packaged. Alternatively, packaging cell lines can be produced by transforming a cell line with one or more expression plasmids encoding the appropriate core and envelope proteins. In these cells, the *gag*, *pol*, and *env* genes can be derived from the same or different retroviruses.

A number of packaging cell lines suitable for the present invention are also available in the prior art. Examples of these cell lines include Crip, GPE86, PA317 and PG13. *See* Miller *et al.*, *J. Virol.* 65:2220-2224 (1991), which is incorporated herein by reference. Examples of other packaging cell lines are described in Cone, R. and Mulligan, R.C., *Proceedings of the National Academy of Sciences, USA,* 81:6349-6353 (1984) and in Danos, O. and R.C. Mulligan, *Proceedings of the National Academy of Sciences, USA,* 85: 6460-6464 (1988), Eglitis, M.A., *et al.* (1988), *supra*, and Miller, A.D., (1990), *supra*, also all incorporated herein by reference.

Packaging cell lines capable of producing retroviral vector particles with chimeric envelope proteins may be used. Alternatively, amphotropic or xenotropic envelope proteins, such as those produced by PA317 and GPX packaging cell lines may be used to package the retroviral vectors.

In some embodiments of the invention, an antisense nucleic acid is administered which hybridizes to an gene associated with cancer development or to

a transcript thereof. The antisense nucleic acid can be provided as an antisense oligonucleotide (*see, e.g.*, Murayama *et al.*, *Antisense Nucleic Acid Drug Dev.* 7:109-114 (1997)). Genes encoding an antisense nucleic acid can also be provided; such genes can be introduced into cells by methods known to those of skill in the art. For example, one can introduce a gene that encodes an antisense nucleic acid in a viral vector, such as, for example, in hepatitis B virus (*see, e.g.*, Ji *et al.*, *J. Viral Hepat.* 4:167-173 (1997)), in adeno-associated virus (*see, e.g.*, Xiao *et al.*, *Brain Res.* 756:76-83 (1997)), or in other systems including, but not limited, to an HVJ (Sendai virus)-liposome gene delivery system (*see, e.g.*, Kaneda *et al.*, *Ann. N.Y. Acad. Sci.* 811:299-308 (1997)), a "peptide vector" (*see, e.g.*, Vidal *et al.*, *CR Acad. Sci III* 32:279-287 (1997)), as a gene in an episomal or plasmid vector (*see, e.g.*, Cooper *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 94:6450-6455 (1997), Yew *et al. Hum Gene Ther.* 8:575-584 (1997)), as a gene in a peptide-DNA aggregate (*see, e.g.*, Niidome *et al.*, *J. Biol. Chem.* 272:15307-15312 (1997)), as "naked DNA" (*see, e.g.*, U.S. 5,580,859 and U.S. 5,589,466), in lipidic vector systems (*see, e.g.*, Lee *et al.*, *Crit Rev Ther Drug Carrier Syst.* 14:173-206 (1997)), polymer coated liposomes (Marin *et al.*, United States Patent No. 5,213,804, issued May 25, 1993; Woodle *et al.*, United States Patent No. 5,013,556, issued May 7, 1991), cationic liposomes (Epand *et al.*, United States Patent No. 5,283,185, issued February 1, 1994; Jessee, J.A., United States Patent No. 5,578,475, issued November 26, 1996; Rose *et al*, United States Patent No. 5,279,833, issued January 18, 1994; Gebeyehu *et al.*, United States Patent No. 5,334,761, issued August 2, 1994), gas filled microspheres (Unger *et al.*, United States Patent No. 5,542,935, issued August 6, 1996), ligand-targeted encapsulated macromolecules (Low *et al.* United States Patent No. 5,108,921, issued April 28, 1992; Curiel *et al.*, United States Patent No. 5,521,291, issued May 28, 1996; Groman *et al.*, United States Patent No. 5,554,386, issued September 10, 1996; Wu *et al.*, United States Patent No. 5,166,320, issued November 24, 1992).

C. Pharmaceutical Formulations

When used for pharmaceutical purposes, the vectors used for gene therapy are formulated in a suitable buffer, which can be any pharmaceutically acceptable buffer, such as phosphate buffered saline or sodium phosphate/sodium

sulfate, Tris buffer, glycine buffer, sterile water, and other buffers known to the ordinarily skilled artisan such as those described by Good *et al.* (1966) *Biochemistry* 5:467.

The compositions can additionally include a stabilizer, enhancer or other pharmaceutically acceptable carriers or vehicles. A pharmaceutically acceptable carrier can contain a physiologically acceptable compound that acts, for example, to stabilize the nucleic acids of the invention and any associated vector. A physiologically acceptable compound can include, for example, carbohydrates, such as glucose, sucrose or dextrans, antioxidants, such as ascorbic acid or glutathione, chelating agents, low molecular weight proteins or other stabilizers or excipients. Other physiologically acceptable compounds include wetting agents, emulsifying agents, dispersing agents or preservatives, which are particularly useful for preventing the growth or action of microorganisms. Various preservatives are well known and include, for example, phenol and ascorbic acid. Examples of carriers, stabilizers or adjuvants can be found in Martin, *Remington's Pharm. Sci.*, 15th Ed. (Mack Publ. Co., Easton, PA 1975), which is incorporated herein by reference.

D. Administration of Formulations

The formulations of the invention can be delivered to any tissue or organ using any delivery method known to the ordinarily skilled artisan for example. In some embodiments of the invention, the nucleic acids of the invention are formulated in mucosal, topical, and/or buccal formulations, particularly mucoadhesive gel and topical gel formulations. Exemplary permeation enhancing compositions, polymer matrices, and mucoadhesive gel preparations for transdermal delivery are disclosed in U.S. Patent No. 5,346,701. In some embodiments of the invention, a therapeutic agent is formulated in ophthalmic formulations for administration to the eye.

E. Methods of Treatment

The gene therapy formulations of the invention are typically administered to a cell. The cell can be provided as part of a tissue, such as an

epithelial membrane, or as an isolated cell, such as in tissue culture. The cell can be provided *in vivo*, *ex vivo*, or *in vitro*.

The formulations can be introduced into the tissue of interest *in vivo* or *ex vivo* by a variety of methods. In some embodiments of the invention, the nucleic acids of the invention are introduced to cells by such methods as microinjection, calcium phosphate precipitation, liposome fusion, or biolistics. In further embodiments, the nucleic acids are taken up directly by the tissue of interest.

In some embodiments of the invention, the nucleic acids of the invention are administered *ex vivo* to cells or tissues explanted from a patient, then returned to the patient. Examples of *ex vivo* administration of therapeutic gene constructs include Arteaga *et al.*, *Cancer Research* 56(5):1098-1103 (1996); Nolta *et al.*, *Proc Natl. Acad. Sci. USA* 93(6):2414-9 (1996); Koc *et al.*, *Seminars in Oncology* 23 (1):46-65 (1996); Raper *et al.*, *Annals of Surgery* 223(2):116-26 (1996); Dalesandro *et al.*, *J. Thorac. Cardi. Surg.*, 11(2):416-22 (1996); and Makarov *et al.*, *Proc. Natl. Acad. Sci. USA* 93(1):402-6 (1996).

VI. General Recombinant Nucleic Acids Methods for Use with the Invention

A. General Recombinant Nucleic Acids Methods

Nucleotide sizes are given in either kilobases (kb) or base pairs (bp). These are estimates derived from agarose or acrylamide gel electrophoresis or, alternatively, from published DNA sequences.

Oligonucleotides that are not commercially available can be chemically synthesized according to the solid phase phosphoramidite triester method first described by S.L. Beaucage and M.H. Caruthers, *Tetrahedron Letts.*, 22(20):1859-1862 (1981), using an automated synthesizer, as described in D.R. Needham Van Devanter *et. al.*, *Nucleic Acids Res.*, 12:6159-6168, 1984. Purification of oligonucleotides is, for example, by either native acrylamide gel electrophoresis or by anion-exchange HPLC as described in J.D. Pearson and F.E. Reanier, *J. Chrom.*, 255:137-149, 1983.

The nucleic acids described here, or fragments thereof, can be used as a hybridization probe for a cDNA library to isolate the corresponding full length

cDNA and to isolate other cDNAs which have a high sequence similarity to the gene or similar biological activity. Probes of this type preferably have at least 30 bases and may contain, for example, 50 or more bases. The probe may also be used to identify a cDNA clone corresponding to a full length transcript and a genomic clone or clones that contain the complete gene, including regulatory and promotor regions, exons and introns. An example of such a screen includes isolating the coding region of the gene by using the known DNA sequence to synthesize an oligonucleotide probe. Labeled oligonucleotides having a sequence complementary to that of the nucleic acids of the present invention can be used to screen a library of human cDNA, genomic DNA or mRNA to determine which members of the library the probe hybridizes to.

The sequence of the cloned genes and synthetic oligonucleotides can be verified using the chemical degradation method of A.M. Maxam *et al.*, *Methods in Enzymology*, 65:499560, (1980). The sequence can be confirmed after the assembly of the oligonucleotide fragments into the double-stranded DNA sequence using the method of Maxam and Gilbert, *supra*, or the chain termination method for sequencing double-stranded templates of R.B. Wallace *et al., Gene*, 16:21-26, 1981. Southern blot hybridization techniques can be carried out according to Southern *et al., J. Mol. Biol.*, 98:503, 1975.

B. Cloning Methods for the Isolation of Nucleotide Sequences Encoding the Desired Proteins

In general, the nucleic acids encoding the subject proteins are cloned from DNA sequence libraries that are made to encode copy DNA (cDNA) or genomic DNA. The particular sequences can be located by hybridizing with an oligonucleotide probe, the sequence of which can be derived from the sequence listing provided herein, which provides a reference for PCR primers and defines suitable regions for isolating cancer-associated specific probes. Alternatively, where the sequence is cloned into an expression library, the expressed recombinant protein can be detected immunologically with antisera or purified antibodies made against the cancer-associated protein of interest.

To make the cDNA library, one should choose a source that is rich in mRNA. The mRNA can then be made into cDNA, ligated into a recombinant

vector, and transfected into a recombinant host for propagation, screening and cloning. Methods for making and screening cDNA libraries are well known. *See*, Gubler, U. and Hoffman, B.J., *Gene* 25:263-269, 1983 and Sambrook, *supra*.

For a genomic library, the DNA is extracted from the tissue and either mechanically sheared or enzymatically digested to yield fragments of preferably about 5-100 kb. The fragments are then separated by gradient centrifugation from undesired sizes and are constructed in bacteriophage lambda vectors. These vectors and phage are packaged *in vitro,* as described in Sambrook, *supra*. Recombinant phages are analyzed by plaque hybridization as described in Benton and Davis, *Science*, 196:180-182 (1977). Colony hybridization is carried out as generally described in M. Grunstein *et al.*, *Proc. Natl. Acad. Sci. USA.*, 72:3961-3965 (1975).

An alternative method combines the use of synthetic oligonucleotide primers with polymerase extension on an mRNA or DNA template. This polymerase chain reaction (PCR) method amplifies the nucleic acids encoding the protein of interest directly from mRNA, cDNA, genomic libraries or cDNA libraries. Restriction endonuclease sites can be incorporated into the primers. Polymerase chain reaction or other *in vitro* amplification methods may also be useful, for example, to clone nucleic acids encoding specific proteins and express said proteins, to synthesize nucleic acids that will be used as probes for detecting the presence of mRNA encoding cancer-associated proteins in physiological samples, for nucleic acid sequencing, or for other purposes. U.S. Patent Nos. 4,683,195 and 4,683,202 describe this method. Genes amplified by a PCR reaction can be purified from agarose gels and cloned into an appropriate vector.

Appropriate primers and probes for identifying the genes encoding cancer-associated proteins from mammalian tissues are generated from comparisons of the sequences provided herein. For a general overview of PCR, *see PCR Protocols: A Guide to Methods and Applications*. (Innis, M, Gelfand, D., Sninsky, J. and White, T., eds.), *Academic Press*, San Diego (1990), incorporated herein by reference.

Synthetic oligonucleotides can be used to construct genes. This is done using a series of overlapping oligonucleotides, usually 40-120 bp in length,

representing both the sense and anti-sense strands of the gene. These DNA fragments are then annealed, ligated and cloned.

A gene involved in the onset of cancer, for example, can be cloned using intermediate vectors before transformation into mammalian cells for expression. These intermediate vectors are typically prokaryote vectors or shuttle vectors. The proteins can be expressed in either prokaryotes, using standard methods well known to those of skill in the art, or eukaryotes as described *infra*.

C. Expression in Eukaryotes

Standard eukaryotic transfection methods are used to produce eukaryotic cell lines, *e.g.*, yeast, insect, or mammalian cell lines, which express large quantities of the cancer-associated proteins which are then purified using standard techniques. *See, e.g.*, Colley *et al.*, *J. Biol. Chem.* 264:17619-17622, (1989), and Guide to Protein Purification, in Vol. 182 of *Methods in Enzymology* (Deutscher ed., 1990), both of which are incorporated herein by reference.

Transformations of eukaryotic cells are performed according to standard techniques as described by D.A. Morrison, *J. Bact.*, 132:349-351 (1977), or by J.E. Clark-Curtiss and R. Curtiss, *Methods in Enzymology*, 101:347-362, Eds. R. Wu *et. al.*, Academic Press, New York (1983).

Any of the well known procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, polybrene, protoplast fusion, electroporation, liposomes, microinjection, plasma vectors, viral vectors and any of the other well known methods for introducing cloned genomic DNA, cDNA, synthetic DNA or other foreign genetic material into a host cell (see Sambrook *et al.*, *supra*). It is only necessary that the particular genetic engineering procedure utilized be capable of successfully introducing at least one gene into the host cell which is capable of expressing the protein.

The particular eukaryotic expression vector used to transport the genetic information into the cell is not particularly critical. Any of the conventional vectors used for expression in eukaryotic cells may be used. Expression vectors containing regulatory elements from eukaryotic viruses are typically used. SV40 vectors include pSVT7 and pMT2. Vectors derived from bovine papilloma virus

include pBV-1MTHA, and vectors derived from Epstein Bar virus include pHEBO, and p2O5. Other exemplary vectors include pMSG, pAV009/A$^+$, pMTO10/A$^+$, pMAMneo-5, baculovirus pDSVE, and any other vector allowing expression of proteins under the direction of the SV-40 early promoter, SV-40 later promoter, metallothionein promoter, murine mammary tumor virus promoter, Rous sarcoma virus promoter, polyhedrin promoter, or other promoters shown effective for expression in eukaryotic cells.

The vectors usually include selectable markers which result in gene amplification, such as thymidine kinase, aminoglycoside phosphotransferase, hygromycin B phosphotransferase, xanthine-guanine phosphoribosyl transferase, CAD (carbamyl phosphate synthetase, aspartate transcarbamylase, and dihydroorotase), adenosine deaminase, dihydrofolate reductase, and asparagine synthetase and ouabain selection. Alternatively, high yield expression systems not involving gene amplification are also suitable, such as using a baculovirus vector in insect cells, with a target protein encoding sequence under the direction of the polyhedrin promoter or other strong baculovirus promoters.

The expression vector of the present invention will typically contain both prokaryotic sequences that facilitate the cloning of the vector in bacteria as well as one or more eukaryotic transcription units that are expressed only in eukaryotic cells, such as mammalian cells. The vector may or may not comprise a eukaryotic replicon. If a eukaryotic replicon is present, then the vector is amplifiable in eukaryotic cells using the appropriate selectable marker. If the vector does not comprise a eukaryotic replicon, no episomal amplification is possible. Instead, the transfected DNA integrates into the genome of the transfected cell, where the promoter directs expression of the desired gene. The expression vector is typically constructed from elements derived from different, well characterized viral or mammalian genes. For a general discussion of the expression of cloned genes in cultured mammalian cells, *see*, Sambrook *et al.*, *supra*, Ch. 16.

The prokaryotic elements that are typically included in the mammalian expression vector include a replicon that functions in *E. coli*, a gene encoding antibiotic resistance to permit selection of bacteria that harbor recombinant plasmids, and unique restriction sites in nonessential regions of the plasmid to allow insertion of eukaryotic sequences. The particular antibiotic resistance gene chosen

is not critical, any of the many resistance genes known in the art are suitable. The prokaryotic sequences are preferably chosen such that they do not interfere with the replication of the DNA in eukaryotic cells.

The expression vector contains a eukaryotic transcription unit or expression cassette that contains all the elements required for the expression of the cancer-associated protein encoding DNA in eukaryotic cells. A typical expression cassette contains a promoter operably linked to the DNA sequence encoding the cancer-associated protein and signals required for efficient polyadenylation of the transcript. The DNA sequence encoding the protein may typically be linked to a cleavable signal peptide sequence to promote secretion of the encoded protein by the transformed cell. Such signal peptides would include, among others, the signal peptides from tissue plasminogen activator, insulin, and neuron growth factor, and juvenile hormone esterase of *Heliothis virescens*. Additional elements of the cassette may include enhancers and, if genomic DNA is used as the structural gene, introns with functional splice donor and acceptor sites.

Eukaryotic promoters typically contain two types of recognition sequences, the TATA box and upstream promoter elements. The TATA box, located 25-30 base pairs upstream of the transcription initiation site, is thought to be involved in directing RNA polymerase to begin RNA synthesis. The other upstream promoter elements determine the rate at which transcription is initiated.

Enhancer elements can stimulate transcription up to 1,000 fold from linked homologous or heterologous promoters. Enhancers are active when placed downstream or upstream from the transcription initiation site. Many enhancer elements derived from viruses have a broad host range and are active in a variety of tissues. For example, the SV40 early gene enhancer is suitable for many cell types. Other enhancer/promoter combinations that are suitable for the present invention include those derived from polyoma virus, human or murine cytomegalovirus, the long term repeat from various retroviruses such as murine leukemia virus, murine or Rous sarcoma virus and HIV. *See, Enhancers and Eukaryotic Expression*, Cold Spring Harbor Pres, Cold Spring Harbor, N.Y. 1983, which is incorporated herein by reference.

In the construction of the expression cassette, the promoter is preferably positioned about the same distance from the heterologous transcription

start site as it is from the transcription start site in its natural setting. As is known in the art, however, some variation in this distance can be accommodated without loss of promoter function.

In addition to a promoter sequence, the expression cassette should also contain a transcription termination region downstream of the structural gene to provide for efficient termination. The termination region may be obtained from the same gene as the promoter sequence or may be obtained from different genes.

If the mRNA encoded by the structural gene is to be efficiently translated, polyadenylation sequences are also commonly added to the vector construct. Two distinct sequence elements are required for accurate and efficient polyadenylation: GU or U rich sequences located downstream from the polyadenylation site and a highly conserved sequence of six nucleotides, AAUAAA, located 11-30 nucleotides upstream. Termination and polyadenylation signals that are suitable for the present invention include those derived from SV40, or a partial genomic copy of a gene already resident on the expression vector.

In addition to the elements already described, the expression vector of the present invention may typically contain other specialized elements intended to increase the level of expression of cloned genes or to facilitate the identification of cells that carry the transfected DNA. For instance, a number of animal viruses contain DNA sequences that promote the extra chromosomal replication of the viral genome in permissive cell types. Plasmids bearing these viral replicons are replicated episomally as long as the appropriate factors are provided by genes either carried on the plasmid or with the genome of the host cell.

The cDNA encoding the protein of the invention can be ligated to various expression vectors for use in transforming host cell cultures. The vectors typically contain gene sequences to initiate transcription and translation of the cancer-associated gene. These sequences need to be compatible with the selected host cell. In addition, the vectors preferably contain a marker to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or metallothionein. Additionally, a vector might contain a replicative origin.

Cells of mammalian origin are illustrative of cell cultures useful for the production of, for example, the cancer-associated protein. Mammalian cell

systems often will be in the form of monolayers of cells although mammalian cell suspensions may also be used. Illustrative examples of mammalian cell lines include VERO and HeLa cells, Chinese hamster ovary (CHO) cell lines, WI38, BHK, COS-7 or MDCK cell lines. NIH 3T3 or COS cells are preferred.

As indicated above, the vector, *e.g.*, a plasmid, which is used to transform the host cell, preferably contains DNA sequences to initiate transcription and sequences to control the translation of the cancer-associated protein gene sequence. These sequences are referred to as expression control sequences. Illustrative expression control sequences are obtained from the SV-40 promoter (*Science*, 222:524-527 (1983)), the CMV I.E. Promoter (*Proc. Natl. Acad. Sci.* 81:659-663 (1984)) or the metallothionein promoter (*Nature* 296:39-42 (1982)). The cloning vector containing the expression control sequences is cleaved using restriction enzymes and adjusted in size as necessary or desirable and ligated with sequences encoding the cancer-associated protein by means well known in the art.

When higher animal host cells are employed, polyadenylation or transcription terminator sequences from known mammalian genes need to be incorporated into the vector. An example of a terminator sequence is the polyadenylation sequence from the bovine growth hormone gene. Sequences for accurate splicing of the transcript may also be included. An example of a splicing sequence is the VP1 intron from SV40 (Sprague, J. *et al.*, *J. Virol.* 45: 773-781, (1983)).

Additionally, gene sequences to control replication in the host cell may be incorporated into the vector such as those found in bovine papilloma virus type-vectors. Saveria-Campo, M., *"Bovine Papilloma virus DNA a Eukaryotic Cloning Vector"* in DNA Cloning Vol.II a Practical Approach Ed. D.M. Glover, IRL Press, Arlington, Virginia pp. 213-238, (1985).

The transformed cells are cultured by means well known in the art. For example, such means are published in *Biochemical Methods in Cell Culture and Virology*, Kuchler, R.J., Dowden, Hutchinson and Ross, Inc. (1977). The expressed protein is isolated from cells grown as suspensions or as monolayers. The latter are recovered by well known mechanical, chemical or enzymatic means.

VII. Purification of the Proteins for Use with the Invention

After expression, the proteins of the present invention can be purified to substantial purity by standard techniques, including selective precipitation with substances as ammonium sulfate, column chromatography, immunopurification methods, and other methods known to those of skill in the art. *See*, for instance, R. Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag: New York (1982), U.S. Patent No. 4,673,641, Ausubel, and Sambrook, incorporated herein by reference.

A number of conventional procedures can be employed when a recombinant protein is being purified. For example, proteins having established molecular adhesion properties can be reversibly fused to the subject protein. With the appropriate ligand, the cancer-associated protein, for example, can be selectively adsorbed to a purification column and then freed from the column in a relatively pure form. The fused protein is then removed by enzymatic activity. Finally, cancer-associated protein can be purified using immunoaffinity columns.

A. Purification of Proteins from Recombinant Bacteria

When recombinant proteins are expressed by the transformed bacteria in large amounts, typically after promoter induction, although expression can be constitutive, the proteins may form insoluble aggregates. There are several protocols that are suitable for purification of protein inclusion bodies. For example, purification of aggregate proteins (hereinafter referred to as inclusion bodies) typically involves the extraction, separation and/or purification of inclusion bodies by disruption of bacterial cells typically, but not limited by, incubation in a buffer of about 100-150 μg/ml lysozyme and 0.1% Nonidet P40, a non-ionic detergent. The cell suspension can be ground using a Polytron grinder (Brinkman Instruments, Westbury, N.Y.). Alternatively, the cells can be sonicated on ice. Alternate methods of lysing bacteria are described in Ausubel and Sambrook and will be apparent to those of skill in the art.

The cell suspension is generally centrifuged and the pellet containing the inclusion bodies resuspended in buffer which does not dissolve but washes the inclusion bodies, *e.g.*, 20 mM Tris-HCl (pH 7.2), 1 mM EDTA, 150 mM NaCl and

2% Triton-X 100, a non-ionic detergent. It may be necessary to repeat the wash step to remove as much cellular debris as possible. The remaining pellet of inclusion bodies may be resuspended in an appropriate buffer (*e.g.,* 20 mM sodium phosphate, pH 6.8, 150 mM NaCl). Other appropriate buffers will be apparent to those of skill in the art.

Following the washing step, the inclusion bodies are solubilized by the addition of a solvent that is both a strong hydrogen acceptor and a strong hydrogen donor (or a combination of solvents each having one of these properties). The proteins that formed the inclusion bodies may then be renatured by dilution or dialysis with a compatible buffer. Suitable solvents include, but are not limited to, urea (from about 4 M to about 8 M), formamide (at least about 80%, volume/volume basis), and guanidine hydrochloride (from about 4 M to about 8 M). Some solvents which are capable of solubilizing aggregate-forming proteins, such as SDS (sodium dodecyl sulfate) and 70% formic acid, are inappropriate for use in this procedure due to the possibility of irreversible denaturation of the proteins, accompanied by a lack of immunogenicity and/or activity. Although guanidine hydrochloride and similar agents are denaturants, this denaturation is not irreversible and renaturation may occur upon removal (by dialysis, for example) or dilution of the denaturant, allowing re-formation of the immunologically and/or biologically active protein of interest. After solubilization, the protein can be separated from other bacterial proteins by standard separation techniques.

Alternatively, it is possible to purify proteins from bacteria periplasm. Where the protein is exported into the periplasm of the bacteria, the periplasmic fraction of the bacteria can be isolated by cold osmotic shock in addition to other methods known to those of skill in the art (*see,* Ausubel, *supra*). To isolate recombinant proteins from the periplasm, the bacterial cells are centrifuged to form a pellet. The pellet is resuspended in a buffer containing 20% sucrose. To lyse the cells, the bacteria are centrifuged and the pellet is resuspended in ice-cold 5 mM $MgSO_4$ and kept in an ice bath for approximately 10 minutes. The cell suspension is centrifuged and the supernatant decanted and saved. The recombinant proteins present in the supernatant can be separated from the host proteins by standard separation techniques well known to those of skill in the art.

B. Standard Protein Separation Techniques For Purifying Proteins

1. Solubility Fractionation

Often as an initial step, and if the protein mixture is complex, an initial salt fractionation can separate many of the unwanted host cell proteins (or proteins derived from the cell culture media) from the recombinant protein of interest. The preferred salt is ammonium sulfate. Ammonium sulfate precipitates proteins by effectively reducing the amount of water in the protein mixture. Proteins then precipitate on the basis of their solubility. The more hydrophobic a protein is, the more likely it is to precipitate at lower ammonium sulfate concentrations. A typical protocol is to add saturated ammonium sulfate to a protein solution so that the resultant ammonium sulfate concentration is between 20-30%. This will precipitate the most hydrophobic of proteins. The precipitate is discarded (unless the protein of interest is hydrophobic) and ammonium sulfate is added to the supernatant to a concentration known to precipitate the protein of interest. The precipitate is then solubilized in buffer and the excess salt removed if necessary, through either dialysis or diafiltration. Other methods that rely on solubility of proteins, such as cold ethanol precipitation, are well known to those of skill in the art and can be used to fractionate complex protein mixtures.

2. Size Differential Filtration

Based on a calculated molecular weight, a protein of greater and lesser size can be isolated using ultrafiltration through membranes of different pore sizes (for example, Amicon or Millipore membranes). As a first step, the protein mixture is ultrafiltered through a membrane with a pore size that has a lower molecular weight cut-off than the molecular weight of the protein of interest. The retentate of the ultrafiltration is then ultrafiltered against a membrane with a molecular cut off greater than the molecular weight of the protein of interest. The recombinant protein will pass through the membrane into the filtrate. The filtrate can then be chromatographed as described below.

3. Column Chromatography

The proteins of interest can also be separated from other proteins on the basis of their size, net surface charge, hydrophobicity and affinity for ligands. In addition, antibodies raised against proteins can be conjugated to column matrices and the proteins immunopurified. All of these methods are well known in the art.

It will be apparent to one of skill that chromatographic techniques can be performed at any scale and using equipment from many different manufacturers (*e.g.*, Pharmacia Biotech).

It is noted that many of the sequences described herein are publicly available in GenBank, which is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences (*Nucleic Acids Research* 1998 Jan 1;26(1):1-7).

All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to those of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may be made thereto without departing from the spirit or scope of the appended claims.

Table 1 below indicates genes by identification in the "GeneName" column that demonstrate change in expression with cancer. "Image CloneID" refers to the IMAGE Consortium library clone identification number. "DataBaseID" is the database that lists the gene to which the clone has the highest homology: GB (GenBank), SP (SwissProt), or UG (UniGene). "Accession Number" is the gene identification number in the database indicated in the "DataBaseID" listing. Where a tissue is indicated in the "3X higher in normal tissue than cancer" column, it means that the gene is expressed at higher levels in normal healthy cells vs. the corresponding cancerous tissues. The cancerous tissues at issue include colon 1, kidney 55, kidney 105, kidney 510, kidney 1010, kidney renal cell carcinoma,

breast, breast lobularcarcinoma, liver and/or prostate as designated on the Table. Where a tissue is indicated in the "3X higher in cancerous tissue than normal" column, it means that the expression of the subject gene is significantly increased in cancerous vs. the corresponding normal healthy tissues. For example, the Cysteine-Rich Protein 2 (LifeSpan Master ID 952) is expressed at significantly higher levels in normal healthy colon cells than in the cancerous colon 1 cells.

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
180	MULTISPANNING MEMBRANE PROTEIN P76	51587	GB	U81006	Colon1,	breast,
217	EPIDERMAL 67-KDA TYPE II KERATIN	196022	GB	M10938	Breast lobularcarcinoma,	
260	EUKARYOTIC TRANSLATION INITIATION FACTOR 5	113597	GB	U49436		Breast lobularcarcinoma, Colon1, Kidney renalcellcarcinoma,
266	EPIDERMAL GROWTH FACTOR RECEPTOR PRECURSOR	120975	GB	X00588		Kidney 510, Kidney 55, Kidney renalcellcarcinoma,
328	HYPOTHETICAL 80.8 KD PROTEIN ZC21.4 IN CHROMOSOME	364444	SP	P34588	Colon1,	breast,
373	ALPHA-1-ACID GLYCOPROTEIN 1 PRECURSOR	428793	GB	X02544	Liver,	
414	A2-CHIMAERIN	565488	GB	Z22641	Colon1, Kidney 510, Kidney 55,	
428	HYPOTHETICAL 48.5 KD PROTEIN	668042	GB	U79241	Colon1,	
526	UNKNOWN	22970				Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,
559	FRUCTOSE-BISPHOSPHATE ALDOLASE A	23831	GB	X07292		Colon1, Kidney 1010, Kidney 105, Kidney 510,

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
613	NFIA1-PROTEIN	26230	GB	Y07690		
849	GLYCOGEN (STARCH) SYNTHASE, MUSCLE	35615	GB	J04501	Colon1,	Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma, Liver,
871	CARBONIC ANHYDRASE I	36232	GB	M33987		
935	ZINC FINGER PROTEIN, COMPLETE CDS, CLONE:RES4-26	39475	GB	U95140		breast, Kidney 510, Kidney 55,
952	CYSTEINE-RICH PROTEIN 2	40234	GB	U36190	Colon1,	
1001	SPARC/OSTEONECTIN	41677	GB	J03040		Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,
1025	GEC-1	42377	GB	AF012920	Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,	
1069	HYPOTHETICAL 70.7 KD PROTEIN F09G8.3 IN CHROMOSOME	43847	SP	P34388		Kidney 510, Kidney 55,
1105	PUTATIVE G PROTEIN-COUPLED RECEPTOR (GPR19)	45231	GB	U64871	Kidney 1010, Kidney 510, Kidney 55,	
1117	MEMBRANE GLYCOPROTEIN M6-A	45393	GB	D49958	Kidney 1010, Kidney 105, Kidney 510, Kidney 55,	

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
1129	DNA-DIRECTED RNA POLYMERASE II 140 KD POLYPEPTIDE	46022	GB	X63563	Colon1, Kidney 1010, Kidney 105,	
1133	CAMP-DEPENDENT 3',5'-CYCLIC PHOSPHODIESTERASE 4B	46315	GB	M97515	Colon1, Kidney 55,	
1263	CAM KINASE I	52629	GB	L41816	Colon1, Kidney 1010, Kidney 510, Kidney 55,	
1303	CARBOXYL TERMINAL LIM DOMAIN PROTEIN (CLIM1)	67724	GB	U90878	Colon1,	Kidney renalcellcarcinoma,
1348	RTP	69879	GB	D87953	Colon1,	breast, Kidney 1010, Kidney 105, Kidney 510, Kidney 55,
1378	INSULIN-LIKE GROWTH FACTOR II PRECURSOR	71160	GB	X07868		breast,
1530	ARGINASE 1	78031	GB	M14502		Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,
1621	KIAA0026	81033	GB	D14812		breast,
1642	METALLOTHIONEIN ISOFORM 2	82154	GB	V00594	Liver,	
1644	HAPTOGLOBIN ALPHA 1S (HPA 1S)	82308	GB	X00637		Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,
1650	APOLIPOPROTEIN(A) PRECURSOR	83000	GB	X05199	Liver,	

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
1671	ALDOLASE B (ALDOB)	83854	GB	X02747	Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,	
1692	ALPHA-2-THIOL PROTEINASE INHIBITOR	85133	GB	K02566	Kidney 105, Kidney renalcellcarcinoma,	
2000	KIAA0243	120107	GB	D87683	Kidney 1010, Kidney 105, Kidney 510, Kidney 55,	
2012	PROBABLE SUCCINYL-COA LIGASE	120291	UG	10444	Colon1, Kidney 55, Kidney renalcellcarcinoma, Prostate,	Liver,
2028	MYELOID ELF-1 LIKE FACTOR (MEF)	120944	GB	U32645	Colon1, Kidney 105, Kidney 55,	
2284	COLLAGEN BINDING PROTEIN 2	131751	GB	D83174		breast, Kidney 1010, Kidney 105, Kidney 510, Kidney 55,
2349	PANCREATIC CARCINOMA MARKER PROTEIN GA733-1 PRECUR	135376	GB	X13425	breast, Kidney 1010, Kidney renalcellcarcinoma,	
2411	CALNEXIN	138464	GB	M94859	Colon1,	
2433	NICOTINAMIDE N-METHYLTRANSFERASE (NNMT)	139478	GB	U08021		Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
2435	ATAXIN-1	139537	GB	X79204	Colon1, Kidney 1010, Kidney 105, Kidney 55, Kidney renalcellcarcinoma,	
2442	SERIN PROTEASE WITH IGF-BINDING MOTIF	139997	GB	Y07921		Kidney 1010, Kidney 105, Kidney 510,
2452	INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 4 PRECU	140272	GB	M62403		Kidney renalcellcarcinoma,
2517	PROCOLLAGEN ALPHA 1(I) CHAIN PRECURSOR	143925	GB	Z74615		breast, Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Liver,
2689	ALPHA-ACTININ 1, CYTOSKELETAL ISOFORM	153375	GB	M95178		breast, Breast lobularcarcinoma, Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,
2703	LUNG AMILORIDE SENSITIVE NA+ CHANNEL PROTEIN	153716	GB	X76180	Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,	
2721	14 KDA BETA-GALACTOSIDE-BINDING LECTIN (LL4)	154344	GB	X15256		Kidney 1010, Kidney 105, Kidney 510, Kidney 55,
2730	VON WILLEBRAND FACTOR PRECURSOR	154475	GB	X04385		Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
2738	HLA-DR ANTIGENS ASSOCIATED INVARIANT CHAIN (P33)	154680	GB	X00497		Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,
2818	ZINC FINGER PROTEIN 42 (MZF-1)	159414	GB	M58297		Liver,
2827	MYELOID CELL-SPECIFIC LEUCINE-RICH GLYCOPROTEIN CD	159946	GB	X13334	Colon1, Kidney 510,	
2854	20-KDA MYOSIN LIGHT CHAIN (MLC-2)	162345	GB	J02854	Colon1, Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,	
2859	IMMUNOGLOBULIN LIGHT CHAIN VARIABLE REGION	162999	GB	L26534	Kidney 510,	breast,
2882	N-METHYL-D-ASPARTATE RECEPTOR (NR1-3)	166245	GB	L13268	Breast lobularcarcinoma, Colon1,	
2901	PUTATIVE P64 CLCP PROTEIN	172013	GB	X87689		
2965	T-LYMPHOCYTE MATURATION-ASSOCIATED PROTEIN	176940	GB	M15800	Colon1, Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,	breast,
3002	KIAA0280	180117	GB	D87470	Colon1, Kidney 1010, Kidney 105, Kidney 510, Kidney 55,	

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
3100	HYPOTHETICAL 38.5 KD PROTEIN IN SUI2-TDH2 INTERGEN	193675	SP	P47085	Colon1, Kidney 1010, Kidney 105, Kidney 510, Kidney 55,	
3272	LEUKOCYTE ELASTASE PRECURSOR	205836	GB	M34379	Colon1, Kidney 1010, Kidney 105, Kidney 510, Kidney 55,	
3439	OTK18	229467	GB	D50419	Colon1, Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,	
3484	DNA LIGASE I	233405	GB	M36067	Colon1,	
3516	CALCIUM-BINDING PROTEIN CAN19	238479	GB	M87068		breast, Breast lobularcarcinoma,
3532	HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN UP2	239112	GB	AF039575	Breast lobularcarcinoma,	Colon1, Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,
3560	DESMOPLAKIN I	240961	GB	M77830		breast,
3586	P300/CBP-ASSOCIATED FACTOR (P/CAF)	243927	GB	U57317	Colon1, Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,	
3614	PROCOLLAGEN ALPHA 1(II) CHAIN PRECURSOR	248162	GB	X06268	Colon1, Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,	

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
3680	CD9 ANTIGEN	253168	GB	M38690	Breast lobularcarcinoma,	
3687	CARBONIC ANHYDRASE III	253662	GB	M29458	Colon1, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,	
3700	G PROTEIN-COUPLED RECEPTOR KINASE GRK4	255333	GB	X97879	Colon1, Kidney 1010, Kidney 105, Liver,	
3846	14-3-3 PROTEIN EPSILON	266106	GB	U54778		
3889	RIT PROTEIN	268332	GB	Y07566		Kidney 1010, Kidney 105, Kidney renalcellcarcinoma,
4038	MULTIFUNCTIONAL PROTEIN ADE2H1	280310	GB	X53793		Colon1, Kidney 1010, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,
4192	HEPARIN-BINDING VASCULAR ENDOTHELIAL GROWTH FACTOR	294236	GB	M27281		breast, Colon1,
4209	CYTOCHROME B561	295787	GB	U06715		Kidney 1010, Kidney 105, Kidney 510, Kidney 55,
4228	KRAB ZINC FINGER ZNF75	296847	GB	S67970	Kidney renalcellcarcinoma,	
4257	MYOSIN LIGHT CHAIN 2	298706	GB	M21812	Kidney 1010, Liver,	
4352	CHLORINE CHANNEL PROTEIN P64	302996	SP	P35526		Kidney 1010, Kidney 105, Kidney 510, Kidney 55,
4407	TTF-I	308100	GB	X83973	Colon1,	

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
4472	DI-N-ACETYLCHITOBIASE	321723	GB	M95767		Breast lobularcarcinoma, Kidney renalcellcarcinoma,
4515	ALPHA-MANNOSIDASE	322865	GB	U68382	Colon1, Kidney renalcellcarcinoma,	
4657	TYPE IV COLLAGEN ALPHA (2) CHAIN.	347332	GB	X05610	Colon1,	
4703	PROTEIN PHOSPHATASE INHIBITOR 2 (PPP1R2)	360595	GB	U68106	Kidney 510, Kidney 55, Kidney renalcellcarcinoma,	
4763	TUBULIN ALPHA-1 CHAIN	362509	GB	K00557		Kidney 1010, Kidney 105, Kidney 510, Kidney 55,
4922	HYPOTHETICAL PROTEIN KIAA0076	418130	GB	D38548	Kidney 1010, Kidney 510, Kidney 55,	
4933	INSULIN RECEPTOR (INSR)	427812	GB	X02160	Colon1, Kidney 1010, Kidney 105, Kidney 510, Kidney 55,	
5122	PTP1C/HCP	471091	GB	X82818	Colon1,	
5207	PROTEASOME SUBUNIT X	486166	GB	D29011		
5330	GDP DISSOCIATION INHIBITOR BETA	489635	GB	Y13286	Colon1,	
5353	LYMPHOCYTE FUNCTION ASSOCIATED ANTIGEN-3 (LFA-3)	490368	GB	Y00636	Colon1, Kidney 1010, Kidney 105, Kidney 55, Kidney renalcellcarcinoma,	

LifeSpan Master ID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
5430	NATURAL KILLER CELLS PROTEIN 4 PRECURSOR	504158	GB	M59807	Colon1,	Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,
5539	BILIARY GLYCOPROTEIN	511562	GB	X16354	Colon1,	
5555	HEME OXYGENASE 1	512374	GB	X06985		
5634	CYTOCHROME C OXIDASE POLYPEPTIDE IV PRECURSOR	530049	GB	X54691	Colon1,	
5674	ACETOACETYL-COENZYME A THIOLASE	530942	GB	D90228		Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,
5682	OLIGOMYCIN SENSITIVITY CONFERRING(OSCP) PROTEIN	531268	GB	D13127	Breast lobularcarcinoma,	
5688	PYRROLINE 5-CARBOXYLATE REDUCTASE	531411	GB	M77836		breast, Kidney 1010, Kidney 105, Kidney 510, Kidney 55,
5794	GLUTATHIONE PEROXIDASE	546504	GB	X15667	Breast lobularcarcinoma,	
5891	TITIN	561694	GB	X90568	Colon1,	
5913	RAS-RELATED PROTEIN (KREV-1)	562931	GB	M22995	Colon1,	breast,
5921	ENDOGENOUS RETROVIRUS CLONE ERV MLN PUTATIVE REVER	563318	GB	U27242	Colon1,	

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
5947	B-CELL ACTIVATION PROTEIN	564503	GB	S53354	Colon1,	
5959	SUBUNIT OF COATOMER COMPLEX	565162	GB	X70476	Colon1, Kidney 105, Kidney renalcellcarcinoma,	breast,
6009	INTERFERON-ALPHA-INDUCIBILE P27	587600	GB	X67325		breast,
6011	RIBOSOMAL PROTEIN S6 KINASE II ALPHA 2	587780	SP	P51812	Colon1, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,	
6079	HYPOTHETICAL 30.8 KD PROTEIN	593061	GB	U79274		breast,
6082	HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, DQ	593221	GB	M17847		
6119	COMPLEMENT FOURTH COMPONENT (C4) GAMMA CHAIN (CODO	595629	GB	S81585		Kidney 1010, Kidney 105, Kidney 510, Kidney 55,
6144	KRAB-ZINC FINGER PROTEIN KZF-1	611253	GB	U67082	Colon1, Kidney 510, Kidney 55,	
6180	P311 HUM	624859	GB	U30521	Colon1, Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,	breast, Breast lobularcarcinoma,

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
6219	LIM DOMAIN PROTEIN, CARDIAC	627939	GB	U49837	Colon1, Kidney renalcellcarcinoma,	
6282	SODIUM CHANNEL PROTEIN, BRAIN I ALPHA SUBUNIT	649192	GB	X65361	Colon1, Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,	
6294	VOLTAGE-GATED POTASSIUM CHANNEL PROTEIN KV1.5	650236	GB	M60451		Colon1, Kidney 510, Kidney 55,
6309	HBD-1 PROTEIN	665086	GB	X92744	Kidney 1010, Kidney 510, Kidney renalcellcarcinoma,	
6344	ENKEPHALINASE	684150	GB	X07166	Colon1, Kidney 105, Kidney 55, Kidney renalcellcarcinoma,	
6378	KIAA0409	713274	GB	AB007869		breast,
6396	MHC CLASS I ANTIGEN HLA-A33 (HLA-A 3303)	724379	GB	U09740	Colon1,	Kidney 105, Kidney 55,
6403	GLYCERALDEHYDE-3-PHOSPHATE DEHYDROGENASE	724709	GB	X01677		breast, Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,
6418	CHLORIDE CHANNEL PROTEIN CLC-KA	725931	GB	Z30643	Kidney renalcellcarcinoma,	
6504	MITOCHONDRIAL DNA	741496	GB	X93334	Colon1, Kidney 1010, Kidney 510, Kidney 105, Kidney 55,	

LifeSpan Master ID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
6586	TRANSCRIPTION INITIATION FACTOR IIE, ALPHA SUBUNIT	760220	GB	X63468	Colon1, Kidney 1010, Kidney 105, Kidney 510, Kidney 55,	
6602	COMPLEMENT C3 PRECURSOR	771021	GB	K02765		Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,
6709	FIBRINOGEN BETA CHAIN	82410	GB	J00129		Kidney 1010, Kidney 105, Kidney 510, Kidney 55,
6710	FIBRINOGEN BETA CHAIN	83050	GB	J00129		breast, Kidney 1010, Kidney 105, Kidney 510, Kidney 55,
6773	CDC37 HOMOLOG	25621	GB	U63131		Colon1,
6864	UNKNOWN	136282				
6874	C1Q B-CHAIN OF COMPLEMENT SYSTEM	141461	GB	X03084		Kidney 1010, Kidney 105, Kidney 510, Kidney 55,
6930	CLONE 23915	44415	GB	AF038197		
6933	MAP KINASE KINASE	44727	GB	X96757	breast,	Liver,
6951	TYROSINE KINASE RECEPTOR (AXL)	49318	GB	M76125	breast, Breast lobularcarcinoma,	Kidney 510, Kidney 55, Kidney renalcellcarcinoma,
7013	UNKNOWN	191592			Colon1, Kidney 1010, Kidney 105, Kidney renalcellcarcinoma,	
7039	ACROSOMAL SERINE PROTEASE INHIBITOR	206458	GB	S58545		
7180	COMPLEMENT C1R COMPONENT PRECURSOR	299611	GB	X04701		Kidney 1010, Kidney 105, Kidney 510, Kidney 55,

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
7188	KIAA0075	304908	GB	D38550		Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,
7208	MYOSIN LIGHT CHAIN KINASE, SMOOTH MUSCLE AND NON-M	310019	GB	X90870	breast, Breast lobularcarcinoma,	
7212	MICROSOMAL GLUTATHIONE S-TRANSFERASE	310414	GB	J03746	Liver,	Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Kidney renalcellcarcinoma,
7300	SERINE PROTEASE WITH IGF-BINDING MOTIF	347396	GB	Y07921		Kidney 1010, Kidney 105, Kidney 510, Kidney 55,
7325	HYPOTHETICAL 80.8 KD PROTEIN ZC21.4 IN CHROMOSOME	364444	SP	P34588	Colon1,	
7345	CYCLIN G1 INTERACTING PROTEIN (1500GX1)	380676	GB	U61835	Colon1, Kidney 105, Kidney 55,	Prostate,
7420	UNKNOWN	502507			Kidney 1010, Kidney 510, Kidney renalcellcarcinoma,	breast, Liver,
7523	KERATIN-RELATED PROTEIN	592758	GB	X05803	Colon1,	breast, Breast lobularcarcinoma,
7568	PYRUVATE KINASE, M1	650222	GB	M26252	Kidney 510, Kidney 55,	Breast lobularcarcinoma,
7640	G1/S-SPECIFIC CYCLIN D2	48206	GB	D13639	Colon1,	
7435	PUTATIVE SURFACE GLYCOPROTEIN PRECURSOR	505491	GB	L48984	Colon1, Kidney 510, Kidney 55,	

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
1806	PLATELET FACTOR 4 PRECURSOR	111349	GB	M25897	Breast lobularcarcinoma, Colon1, Kidney 1010, Kidney 510,	
4626	MYOSIN VI (MYO6)	338751	GB	U90236	Prostate,	
3555	C4B-BINDING PROTEIN BETA CHAIN PRECURSOR	240712	GB	L11245	Liver,	
6382	MAGO NASHI PROTEIN HOMOLOG	713812	SP	P50606	breast,	
5067	15-HYDROXYPROSTAGLAND IN DEHYDROGENASE	469772	GB	L76465		
213	MICROTUBULE-ASSOCIATED PROTEIN 1A (MAP1A)	180933	GB	U38292	Breast lobularcarcinoma, Colon1, Prostate,	Kidney 1010,
5104	BILIVERDIN REDUCTASE A PRECURSOR	470598	GB	U34877	Breast lobularcarcinoma, Kidney 1010,	
4296	TRYPSINOGEN, ANIONIC PRECURSOR	300611	UG	50915	Kidney 105, Kidney 55,	
4792	TRANSCRIPTION FACTOR AP-2 BETA	363520	GB	X95694	breast,	
3252	PUTATIVE IL-16 PROTEIN PRECURSOR,	204381	GB	M90391	Colon1, Kidney renalcellcarcinoma,	
6093	TRANSCRIPTION FACTOR ERF-1	594372	GB	U85658	Prostate,	
1706	ANGIOGENIN PRECURSOR	85966	GB	M11567		

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
2008	CARBONIC ANHYDRASE II	120247	GB	Y00339		
3983	PAPS SYNTHETASE	276543	GB	Y10387	Colon1,	Breast lobularcarcinoma,
3991	KIAA0185	277047	GB	D80007	Breast lobularcarcinoma, Colon1, Kidney renalcellcarcinoma, Liver,	
215	RED CELL ANION EXCHANGER (EPB3, AE1, BAND 3)	195416	GB	X77738		
812	CHLORINE CHANNEL PROTEIN (P64)	34260	GB	L16547	breast, Kidney renalcellcarcinoma,	
364	SEX HORMONE-BINDING GLOBULIN PRECURSOR	417967	GB	X05885	Breast lobularcarcinoma, Prostate,	
3790	ZINC-FINGER PROTEIN HT2A	262223	GB	U18543	breast, Colon1, Kidney 1010, Kidney 510, Kidney 55,	
3343	AMILORIDE-SENSITIVE SODIUM CHANNEL ALPHA-SUBUNIT	212704	GB	D79992	Colon1, Kidney renalcellcarcinoma,	
537	GAMMA-AMINOBUTYRIC ACID TYPE A RECEPTOR ALPHA 6 SU	23353	GB	S81944	breast, Kidney 105,	
7092	UNKNOWN	259390			breast, Breast lobularcarcinoma, Colon1, Kidney 1010, Kidney 105, Prostate,	

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
4491	UNKNOWN	322334			Colon1,	Breast lobularcarcinoma,
4572	UNKNOWN	324945			breast, Breast lobularcarcinoma, Kidney 1010, Kidney 510, Kidney 55, Prostate,	
2054	NEUTROPHIL OXIDASE FACTOR (P67-PHOX)	121975	GB	M32011	breast, Breast lobularcarcinoma, Kidney 1010, Kidney 105, Prostate,	
2610	STEROID SULFATASE (MICROSOMAL)	149741	GB	J04964	breast, Colon1,	
308	ADENYLATE CYCLASE, TYPE I	344141	GB	L05500	breast, Breast lobularcarcinoma, Kidney 510, Kidney 55, Prostate,	
4984	GUANINE NUCLEOTIDE EXCHANGE FACTOR PROTEIN TRIO	429234	GB	U42390	Breast lobularcarcinoma, Kidney 510, Kidney 55,	
1306	EARLY GROWTH RESPONSE PROTEIN 1 (HEGR1)	68041	GB	X52541		
6049	ELONGIN A	591460	GB	L47345	Colon1, Kidney 510, Kidney 55,	
389	OXIDOREDUCTASE (HHCMA56)	501939	GB	U13395	Kidney 510, Kidney 55,	
1830	MELTRIN-L PRECURSOR (ADAM12)	112163	GB	AF023477	breast, Breast lobularcarcinoma,	Prostate,

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
1196	UBIQUITIN CARBOXYL-TERMINAL HYDROLASE T	49297	GB	X91349	Kidney 1010, Kidney 510, Liver,	
4384	ZINC FINGER Y-CHROMOSOMAL PROTEIN 1	306222	UG	22879	Colon1,	
210	MIZ-1 PROTEIN	179992	GB	Y09723	Breast lobularcarcinoma, Colon1,	
2121	FICOLIN	124913	GB	D83920	breast, Breast lobularcarcinoma, Kidney 510, Kidney 55, Prostate,	
3975	TRANSLIN ASSOCIATED PROTEIN X	275254	GB	X95073	breast, Breast lobularcarcinoma,	
6446	CYCLIC-AMP-DEPENDENT TRANSCRIPTION FACTOR ATF-1	727546	GB	X55544	Colon1,	
4973	CANALICULAR MULTISPECIFIC ORGANIC ANION TRANSPORTE	428916	GB	U49248	Breast lobularcarcinoma,	
1828	KIAA0156	112084	GB	D63879	breast, Breast lobularcarcinoma,	
3114	UNKNOWN	194250			Breast lobularcarcinoma, Colon1, Prostate,	
1210	MELANOMA-ASSOCIATED CHONDROITIN SULFATE PROTEOGLYC	50211	GB	X96753		

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
3738	C-REL PROTO-ONCOGENE PROTEIN	258589	GB	X75042	Breast lobularcarcinoma, Prostate,	
4986	ZINC FINGER PROTEIN HZF4	429284	GB	X78927	Breast lobularcarcinoma, Kidney 1010, Kidney 510, Kidney 55, Prostate,	
1211	CD86 ANTIGEN	50214	GB	U04343	breast, Breast lobularcarcinoma, Kidney 1010, Kidney 510, Kidney 55, Prostate,	
7406	T12D8.I	489983	UG	108396		Colon1,
3076	ADDUCIN GAMMA SUBUNIT	190822	GB	U37122		Colon1, Kidney 55, Kidney renalcellcarcinoma,
4428	SHC TRANSFORMING PROTEIN	309796	GB	X68148		breast, Kidney 1010, Kidney 105, Kidney 510, Kidney 55,
5498	EUKARYOTIC TRANSLATION INITIATION FACTOR (EIF3)	510245	GB	U78525		breast, Breast lobularcarcinoma, Kidney 1010, Kidney 105, Liver, Prostate,
6892	IROQUOIS-CLASS HOMEODOMAIN PROTEIN IRX-2A	152453	GB	U90304		breast, Breast lobularcarcinoma,
6069	DNA POLYMERASE DELTA SMALL SUBUNIT	592659	GB	U21090		
1566	THYMOSIN BETA-10	79272	GB	M20259		Kidney 1010, Kidney 105, Kidney 510, Kidney 55,

LifeSpan Master ID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
6453	GLUTAMINYL-TRNA SYNTHETASE	727844	GB	X54326		breast, Kidney 105, Kidney 55,
1886	PUTATIVE MITOCHONDRIAL CARRIER YBR291C	114427	UG	7994		Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Liver,
3887	THYROID RECEPTOR INTERACTING PROTEIN 7	268208	GB	L40357		Colon1, Kidney 1010, Kidney 105, Kidney 510, Kidney 55,
868	C2 DOMAIN CONTAINING PI3-KINASE	36126	GB	AJ000008		Colon1,
1191	HFAT PROTEIN	48855	GB	X87241		Colon1, Liver,
5668	SELENOPROTEIN P PRECURSOR	530814	GB	Z11793		Breast lobularcarcinoma,
502	KIAA0253	22490	GB	D87442		breast, Breast lobularcarcinoma,
4487	UNKNOWN	322311			Colon1,	breast, Kidney 1010, Kidney 105,
7444	UBIQUITIN-CONJUGATING ENZYME E2-18 KD	511757	GB	X96427	Prostate,	Colon1, Kidney renalcellcarcinoma,
6882	ENDOGENOUS RETROVIRAL PROTEASE	144767	GB	M27826		Colon1,
2523	GLYCERATE DEHYDROGENASE	144825	SP	P36234	Kidney 510,	Colon1, Kidney 105,
6385	MESOTHELIN OR CAK1 ANTIGEN PRECURSOR	714080	GB	U40434		breast,

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
6875	NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5	142144	UG	92918		Colon1,
5372	ANTIGEN PEPTIDE TRANSPORTER 1	501605	GB	X57522		Breast lobularcarcinoma, Colon1,
6984	UNKNOWN	173389				
6982	UNKNOWN	172567				Breast lobularcarcinoma, Colon1,
202	PROTEIN KINASE C GAMMA (PARTIAL)	167032	GB	Z15114		breast, Liver, Prostate,
6796	UNKNOWN	32320				Colon1,
6599	TRANSCRIPTION FACTOR SP2	770397	GB	D28588		Breast lobularcarcinoma, Kidney renalcellcarcinoma, Liver,
7301	GLUCOSE TRANSPORTER	347397	GB	K03195		breast, Kidney 1010, Kidney 105, Kidney 510, Kidney 55,
6627	PROSTACYCLIN RECEPTOR	774146	GB	D29634		Breast lobularcarcinoma, Colon1, Kidney renalcellcarcinoma,
5006	80K-H PROTEIN (KINASE C SUBSTRATE)	429669	GB	J03075	Colon1,	Kidney renalcellcarcinoma,
1463	78 KD GLUCOSE REGULATED PROTEIN PRECURSOR	74314	GB	X87949		
4304	BETA-2-MICROGLOBULIN	300966	GB	M17987		Breast lobularcarcinoma,

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
3918	ENDOGENOUS RETROVIRUS TYPE C ONCOVIRUS	270385	GB	M74509		Kidney 510,
5038	KIAA0120	469143	GB	D21261		breast,
4315	TENSIN GENE	301357	GB	Z18529		breast, Colon1,
6273	MACMARCKS	648263	GB	X70326		
179	UNKNOWN	47950				Colon1,
2765	VASOPRESSIN V1A RECEPTOR	155723	GB	S73899		breast, Breast lobularcarcinoma, Kidney renalcellcarcinoma,
3112	PEROXISOMAL MEMBRANE PROTEIN 69 (PMP69)	194139	GB	AF009746		Kidney 55,
2663	ATRIAL NATRIURETIC PEPTIDE RECEPTOR A PRECURSOR	152523	GB	X15357		breast,
6716	CALPAIN 1, LARGE	84298	GB	X04366		breast, Breast lobularcarcinoma,
712	KIAA0194	29851	GB	D83778		
4399	PROCOLLAGEN ALPHA 1(I) CHAIN PRECURSOR	307752	GB	Z74615		breast, Breast lobularcarcinoma,
7305	CHIMERA	356704				breast, Kidney 1010, Kidney 105, Kidney 510, Kidney 55,

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
4623	PANCREATIC ELASTASE IIA	338700	GB	M16652	breast, Breast lobularcarcinoma, Kidney 1010, Kidney 105, Kidney 510, Kidney 55, Prostate,	Kidney renalcellcarcinoma,
4309	HETEROGENEOUS NUCLEAR RIBONUCLEOPROTEIN F	301063	GB	L28010	Breast lobularcarcinoma, Colon1, Kidney 1010, Kidney 105,	Prostate,
1815	SERUM PARAOXONASE/ARYLEST ERASE 1	111520	GB	U53784	breast, Breast lobularcarcinoma, Kidney 510, Kidney 55, Prostate,	
2992	ADENOMATOUS POLYPOSIS COLI PROTEIN	179130	SP	P25054	breast, Breast lobularcarcinoma, Prostate,	
280	COLLAGENASE 3	285780	GB	X75308	breast, Breast lobularcarcinoma, Colon1,	
4197	UNKNOWN	294969			breast, Breast lobularcarcinoma, Prostate,	
1182	ATP-BINDING CASSETTE TRANSPORTER 1	48518	SP	P41233	Breast lobularcarcinoma, Colon1, Kidney 510, Prostate,	
4102	SYT	288695	GB	X79201	Colon1, Kidney 105, Kidney 55, Kidney renalcellcarcinoma,	

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
689	P619	28896	GB	U50078	breast, Breast lobularcarcinoma, Kidney 510, Prostate,	
612	C44C1.2 GENE PRODUCT	26429	UG	10463	breast, Breast lobularcarcinoma, Colon1,	Kidney 1010, Kidney 105,
4165	HYPOTHETICAL PROTEIN ZC84.3	291928	UG	108614	Breast lobularcarcinoma, Colon1, Prostate,	
6822	KIAA0349	40394	GB	AB002347	breast, Breast lobularcarcinoma, Kidney 510, Kidney 55, Prostate,	
636	DELTA-CATENIN	26948	GB	U96136	breast, Breast lobularcarcinoma, Kidney 510, Kidney 55, Prostate,	Kidney renalcellcarcinoma,
5008	UNKNOWN	429676			breast, Breast lobularcarcinoma, Prostate,	
2034	1,4-ALPHA-GLUCAN BRANCHING ENZYME (HGBE)	121163	GB	L07956	breast, Breast lobularcarcinoma, Prostate,	
1455	L-ARGININE:GLYCINE AMIDINOTRANSFERASE	74021	GB	X86401	breast, Breast lobularcarcinoma, Kidney renalcellcarcinoma,	Prostate,
4985	PROTEIN REGULATING CYTOKINESIS 1 (PRC1)	429246	GB	AF027514	breast, Breast lobularcarcinoma, Kidney 510, Kidney 55, Prostate,	

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
4920	MICROSOMAL TRIGLYCERIDE TRANSFER PROTEIN	418090	GB	X91148	breast, Breast lobularcarcinoma, Kidney 510, Prostate,	
6337	INTERFERON ALPHA INDUCED TRANSCRIPTIONAL ACTIVATOR	682770	GB	M97934	breast, Breast lobularcarcinoma, Colon1,	Prostate,
3640	CLONE X874	249614	GB	Z47045	Colon1, Kidney 1010, Kidney 105, Kidney renalcellcarcinoma,	Breast lobularcarcinoma,
7551	TYPE IIA MYOSIN HEAVY CHAIN	628195	GB	S73840	breast, Breast lobularcarcinoma, Prostate,	
6263	KRUPPEL-ASSOCIATED BOX (KRAB)	647076	GB	M67509	Colon1, Kidney renalcellcarcinoma, Liver,	
1064	MYELODYSPLASIA/MYELOID LEUKEMIA FACTOR 2 (MLF2)	43597	GB	U57342	breast, Breast lobularcarcinoma, Prostate,	Kidney renalcellcarcinoma,
6253	HUR RNA BINDING PROTEIN	645216	GB	U38175	Breast lobularcarcinoma, Colon1, Kidney renalcellcarcinoma,	
6988	UNKNOWN	174234			breast, Breast lobularcarcinoma, Prostate,	

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
7540	CASPASE-9 PRECURSOR	612577	GB	U56390	Colon1, Kidney renalcellcarcinoma, Prostate,	
4799	RAS-RELATED YPT1 PROTEIN	363872	GB	Y00094		breast, Breast lobularcarcinoma, Colon1, Kidney renalcellcarcinoma,
171	UNKNOWN	40160				breast, Breast lobularcarcinoma, Colon1, Liver,
3526	LDL-PHOSPHOLIPASE A2	238821	GB	U24577	Kidney 1010, Kidney 105, Kidney 510, Kidney 55,	breast, Breast lobularcarcinoma, Colon1, Prostate,
1086	OSTEOPONTIN	44313	GB	J04765		breast, Breast lobularcarcinoma, Liver, Prostate,
7424	LYSOSOMAL-ASSOCIATED MULTITRANSMEMBRANE PROTEIN (L	503209	GB	U51240		breast, Kidney 1010, Kidney 105, Kidney renalcellcarcinoma,
7159	CASEIN KINASE I, DELTA ISOFORM	284450	GB	U29171		breast, Breast lobularcarcinoma, Kidney 1010, Kidney 105, Kidney 55,
4335	INNER MITOCHONDRIAL MEMBRANE TRANSLOCASE TIM23	302212	GB	AF030162	Kidney 1010, Kidney 105,	breast, Breast lobularcarcinoma, Colon1,

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
2229	MITOTIC FEEDBACK CONTROL PROTEIN MADP2 HOMOLOG	129140	GB	U31278		breast, Breast lobularcarcinoma, Colon1,
2675	TISSUE INHIBITOR OF METALLOPROTEINASES-3 (TIMP-3)	152782	GB	S78453		Breast lobularcarcinoma, Colon1, Kidney renalcellcarcinoma,
5901	SARCOPLASMIC RETICULUM HISTIDINE-RICH CALCIUM-BIND	562365	GB	M60052		breast, Breast lobularcarcinoma, Liver,
3080	SERINE/THREONINE PROTEIN KINASE	190924	GB	AF004849		Colon1, Kidney 105, Kidney 510, Kidney 55, Prostate,
7396	PROTEIN KINASE DYRK2	488243	GB	Y13493	Colon1, Kidney 1010, Kidney 105, Kidney 510, Kidney 55,	breast, Breast lobularcarcinoma, Prostate,
174	GLUCOSE-6-PHOSPHATE DEHYDROGENASE (G6PD)	41559	GB	M35604		breast, Kidney renalcellcarcinoma, Liver,
7210	RETINOIC ACID- AND INTERFERON-INDUCIBLE 58K PROTEI	310105	GB	U34605	Kidney 510, Kidney 55,	Breast lobularcarcinoma, Kidney renalcellcarcinoma, Prostate,
4175	RAN GTPASE ACTIVATING PROTEIN 1	292619	GB	X82260		breast, Colon1, Liver,
5339	3-OXOACYL-[ACYL-CARRIER-PROTEIN] SYNTHASE II	489809	SP	P39435	Prostate,	breast, Breast lobularcarcinoma, Kidney 1010, Kidney 105, Kidney 510, Kidney 55,

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
170	INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 5 (IGFB	38481	GB	L27560		Breast lobularcarcinoma, Colon1, Kidney renalcellcarcinoma,
6754	UNKNOWN	120513				breast, Breast lobularcarcinoma, Colon1,
7505	UNKNOWN	588579				breast, Kidney 105, Liver,
169	C-KIT PROTO-ONCOGENE	37621	GB	X06182		breast, Colon1,
185	SERINE/THREONINE PROTEIN KINASE MO15	127719	GB	Y13120	Colon1,	
203	DELAYED-RECTIFIER K+ CHANNEL ALPHA SUBUNIT (KCNS1)	171627	GB	AF043473	Colon1,	
216	TYROSINE-PROTEIN KINASE LYN	196007	GB	M16038	Breast lobularcarcinoma,	
245	TYROSINE KINASE (HTK)	75009	GB	U07695	Colon1,	
310	TYROSINE KINASE (HTK)	357551	GB	U07695	Kidney 105, Kidney 55,	
520	3',5'-CYCLIC AMP PHOSPHODIESTERASE (HSPDE4C1)	22730	GB	Z46632		
597	RYANODINE RECEPTOR	25322	GB	J05200	Kidney renalcellcarcinoma,	
625	ENDOTHELIAL DIFFERENTIATION PROTEIN (EDG-1)	26418	GB	M31210		Kidney renalcellcarcinoma,

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
717	AMILORIDE-SENSITIVE BRAIN SODIUM CHANNEL BNAC1	29692	GB	U57352		
732	2',3'-CYCLIC NUCLEOTIDE 3'-PHOSPHODIESTERASE	30758	GB	M19650		Liver,
751	G PROTEIN-COUPLED RECEPTOR V28	31466	GB	U20350	Breast lobularcarcinoma,	Colon1,
770	PHOSPHORYLASE KINASE GAMMA 2 (PHKG2)	32407	GB	S73483	Breast lobularcarcinoma,	
796	POTASSIUM CHANNEL ALPHA SUBUNIT (KV9.2)	33144	GB	AF008574	Liver,	Colon1,
813	MINERALOCORTICOID RECEPTOR	34262	GB	M16801		Liver,
859	PLATELET-DERIVED GROWTH FACTOR RECEPTOR ALPHA (PDG	35885	GB	M21574	Colon1,	
863	CALCIUM/CALMODULIN-DEPENDENT PROTEIN KINASE TYPE I	36153	SP	P11730	Kidney 1010, Kidney 105,	Liver,
875	NEUROENDOCRINE/BETA -CELL-TYPE CALCIUM CHANNEL ALPH	36581	GB	M83566	Colon1, Kidney 510, Kidney 55,	
920	CLC-7 CHLORIDE CHANNEL PROTEIN	38871	GB	Z67743		breast,
958	5-HYDROXYTRYPTAMINE 2A RECEPTOR	40580	GB	X57830	Breast lobularcarcinoma, Colon1,	

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
960	MITOGEN INDUCED NUCLEAR ORPHAN RECEPTOR (MINOR)	40831	GB	U12767	Breast lobularcarcinoma, Colon1,	
1006	P2X PURINOCEPTOR 4	42118	GB	Y07684		
1008	GLUTAMATE RECEPTOR 2 (HBGR2)	42242	GB	L20814		
1014	RECEPTOR PROTEIN-TYROSINE KINASE TKT PRECURSOR	42188	GB	X74764		Kidney renalcellcarcinoma, Prostate,
1022	PUTATIVE G PROTEIN-COUPLED RECEPTOR (GPR22)	42685	GB	U66581		Breast lobularcarcinoma,
1112	PROTEIN SERINE/THREONINE KINASE ERK1	45089	GB	X60188	Breast lobularcarcinoma,	
1114	DUAL SPECIFICITY MITOGEN-ACTIVATED PROTEIN KINASE	45641	GB	U66839	Breast lobularcarcinoma,	Colon1,
1123	VOLTAGE-GATED CHLORIDE CHANNEL	45966	GB	S77770	Kidney renalcellcarcinoma,	
1125	ADENOSINE A2A RECEPTOR	45788	GB	X68486		
1137	PROTEIN TYROSINE KINASE T-ROR1 (ROR1)	46342	GB	M97675		
1141	CB1 CANNABINOID RECEPTOR (CNR1)	46807	GB	U73304	Breast lobularcarcinoma,	

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
1149	PIM-2 PROTOONCOGENE HOMOLOG PIM-2H	46959	GB	U77735		
1150	GABA-A RECEPTOR, ALPHA 1 SUBUNIT	47230	GB	X14766		
1174	STE20-LIKE KINASE 3 (MST-3)	48121	GB	AF024636		breast,
1195	PUTATIVE NOVEL RECEPTOR KINASE (GPRKG)	49383	GB	U00686		
1242	P2Y PURINOCEPTOR 9	51646	GB	U90322		Liver,
1277	EMR1 HORMONE RECEPTOR	66507	GB	X81479		
1281	PROTEIN KINASE C, BETA-I TYPE	66607	GB	X04795	breast, Breast lobularcarcinoma,	
1283	MHC CLASS I PROMOTER BINDING PROTEIN	66621	GB	X65463	Kidney renalcellcarcinoma,	
1316	CAMP PHOSPHODIESTERASE	68340	GB	L12052		Colon1,
1364	C3A ANAPHYLATOXIN RECEPTOR	70592	GB	Z73157		
1399	CHLORIDE CONDUCTANCE REGULATORY PROTEIN ICLN	72050	GB	X91788		Prostate,
1438	SERINE/THREONINE PROTEIN KINASE	73547	GB	Z34524		Colon1,

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
1569	LEUCINE ZIPPER BEARING KINASE	79406	GB	AB001872		
1705	RECEPTOR TYROSINE KINASE EPH (PARTIAL)	85953	GB	Z27409		Kidney 1010, Kidney 105,
1770	FIBROBLAST GROWTH FACTOR RECEPTOR 4 (FGFR-4)	110324	GB	X57205	Colon1,	
1825	FIBRILLARIN	111988	GB	X56597	Breast lobularcarcinoma,	Prostate,
1984	G PROTEIN-COUPLED RECEPTOR (GPR4)	119199	GB	U21051	Colon1,	
1997	STRESS RESPONSIVE SERINE/THREONINE PROTEIN KINASE	120015	GB	U60207	Colon1,	
2009	KIAA0151	120254	GB	D63485	Kidney 1010, Prostate,	
2011	CHLORIDE CHANNEL PROTEIN 6	120287	GB	X83378		Colon1,
2074	KIAA0561 PROTEIN	123267	GB	AB011133	Colon1, Kidney 510,	
2083	PROBABLE G PROTEIN-COUPLED RECEPTOR HM74	123666	GB	D10923	Kidney renalcellcarcinoma,	
2180	5-HYDROXYTRYPTAMINE 3 RECEPTOR PRECURSOR	127204	GB	S82612		Kidney 105,
2351	SERINE/THREONINE PROTEIN KINASE PCTAIRE-1	135550	GB	X66363		breast,

LifeSpan Master ID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
2487	P2U NUCLEOTIDE RECEPTOR	141852	GB	U07225	Breast lobularcarcinoma,	
2504	ANGIOTENSIN II RECEPTOR	143073	GB	X65699		
2620	INSULIN-LIKE GROWTH FACTOR I RECEPTOR	150361	GB	X04434	Breast lobularcarcinoma,	
2622	PROTEIN TYROSINE PHOSPHATASE CL 100	150423	GB	X68277	breast,	
2634	PROSTANOID FP RECEPTOR	151011	GB	L24470		
2729	BASIC FIBROBLAST GROWTH FACTOR RECEPTOR 1 PRECURSO	154472	GB	X66945		Colon1,
2758	RETINOIC ACID RECEPTOR	155460	GB	X06614	breast, Breast lobularcarcinoma,	
2807	MYOSIN LIGHT CHAIN KINASE, SMOOTH MUSCLE AND NON-M	158426	GB	X90870	breast, Breast lobularcarcinoma,	
2860	AMILORIDE-SENSITIVE SODIUM CHANNEL BETA-SUBUNIT	163045	GB	X87159	Colon1,	
2867	N-METHYL-D-ASPARTATE RECEPTOR SUBUNIT (GRIN1)	163879	GB	U08106	breast, Colon1,	
2947	INWARD RECTIFIER POTASSIUM CHANNEL 2	174801	GB	U24056		

LifeSpan Master ID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
2954	VOLTAGE-GATED POTASSIUM CHANNEL PROTEIN KV1.3	175489	GB	L02750		
2974	CALCIUM, CALMODULIN-DEPENDENT PROTEIN KINASE II BE	177822	GB	U50358	Colon1,	breast,
2996	POTASSIUM CHANNEL (KCNQ2)	179534	GB	AF033348	Kidney renalcellcarcinoma,	
3010	DUFFY BLOOD GROUP ANTIGEN (FYA-B+)	181704	GB	U01839		Kidney renalcellcarcinoma,
3159	PURINERGIC RECEPTOR P2Y5	196488	GB	AF000546		Colon1,
3230	MET PROTO-ONCOGENE	202615	GB	X54559	Liver,	
3394	CGMP PHOSPHODIESTERASE GAMMA-SUBUNIT (PDEG)	219980	GB	M36476		Colon1,
3429	PERIPHERAL PLASMA MEMBRANE PROTEIN CASK	223193	GB	AF035582	Breast lobularcarcinoma, Kidney 510, Prostate,	
3469	SERINE/THREONINE PROTEIN KINASE PLSTIRE	231497	GB	X66365		Kidney renalcellcarcinoma,
3477	SRC-LIKE KINASE (SLK)	232949	GB	M14676	Colon1,	
3559	TGF-BETAIIR ALPHA	240950	GB	D50683		
3736	PHOSPHODIESTERASE	258516	GB	L20966		breast,

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
3777	INWARDLY RECTIFYING POTASSIUM CHANNEL KIR7.1	260707	GB	AJ006128	Colon1,	
3852	PROSTAGLANDIN E2 RECEPTOR EP2 SUBTYPE	266525	GB	L28175		Liver,
3854	PORIN (POR)	266793	GB	L08666	Colon1,	
3892	ROLIPRAM-SENSITIVE 3',5'-CYCLIC AMP PHOSPHODIESTER	268455	GB	U02882		Liver,
3980	CELL DIVISION PROTEIN KINASE 2 (CDK2)	276282	GB	X62071	Colon1, Liver,	
4041	5-HYDROXYTRYPTAMINE 2C RECEPTOR	280371	GB	U49516		
4099	3',5' CYCLIC NUCLEOTIDE PHOSPHODIESTERASE (HSPDE1C	287705	GB	U40372		Liver,
4131	CAMP-SPECIFIC PHOSPHODIESTERASE 8A (PDE8A)	289972	GB	AF056490		Kidney 1010,
4154	T-CELL-SPECIFIC TYROSINE KINASE EMT	291177	GB	S65186	Kidney renalcellcarcinoma,	Liver,
4286	DIHYDROPYRIDINE-SENSITIVE L-TYPE, CALCIUM CHANNEL	300048	GB	M92301		

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
4289	AMP-ACTIVATED PROTEIN KINASE BETA 2 SUBUNIT.	300137	GB	AJ224538		Breast lobularcarcinoma,
4294	ERBB-2 RECEPTOR PROTEIN-TYROSINE KINASE PRECURSOR	300383	GB	X03363	Colon1, Kidney 510,	
4333	SER-THR PROTEIN KINASE PK428	302177	GB	U59305	Colon1,	
4375	HBRM	306033	GB	X72889	Colon1,	
4422	STRESS ACTIVATED PROTEIN KINASE-3	309482	GB	Y10487	Colon1, Prostate,	Kidney 510, Kidney 55,
4673	RETINOIC ACID RECEPTOR RXR-GAMMA	358433	GB	U38480	Prostate,	
4689	K+ CHANNEL BETA SUBUNIT	360213	GB	L39833	breast, Breast lobularcarcinoma,	
4704	RHODOPSIN	360598	GB	U49742		Prostate,
4749	ROD PHOTORECEPTOR CGMP-GATED CHANNEL CNCG)	362167	GB	S42457		
4750	ENDOTHELIN ET-B RECEPTO	362177	GB	S57283		
4758	SERINE/THREONINE PROTEIN KINASE SGK	362359	GB	Y10032	Kidney 105,	
4774	5'-AMP-ACTIVATED PROTEIN KINASE, BETA SUBUNIT	362744	GB	Y12556		

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
4779	TESTIS-SPECIFIC CAMP-DEPENDENT PROTEIN KINASE CATA	362926	GB	M34181		
4834	CASEIN KINASE I, ALPHA ISOFORM	381589	GB	U59166		
4837	GLUTAMATE RECEPTOR 7 PRECURSOR	381812	GB	U16127		
4838	POTASSIUM CHANNEL KV2.1	381974	GB	L02840		Breast lobularcarcinoma,
5037	CALMODULIN-DEPENDENT PROTEIN KINASE II-DELTA DASH	430337	GB	D14906		
5119	FIBROBLAST GROWTH FACTOR RECEPTOR (K-SAM)	470965	GB	M87770	Colon1,	
5163	DIHYDROPYRIDINE-SENSITIVE L-TYPE, CALCIUM CHANNEL	484731	GB	U07139		
5188	THROMBOXANE A2 RECEPTOR	485744	GB	D38081		
5358	CALCIUM CHANNEL L-TYPE ALPHA 1 SUBUNIT (CACNL1A1)	491064	GB	L29536	Colon1,	
5362	CASEIN KINASE I EPSILON	491245	GB	L37043	Kidney 105, Kidney renalcellcarcinoma,	

LifeSpan Master ID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
5410	PROSTAGLANDIN E RECEPTOR (EP3A1)	503146	GB	X83857	breast, Colon1, Kidney 510,	Liver,
5425	ENDOTHELIN-1 RECEPTOR.	504085	GB	X61950		
5652	TYPE I REGULATORY SUBUNIT OF CAMP-DEPENDENT PROTEI	530348	GB	M17086		breast, Breast lobularcarcinoma,
5928	GABA-A RECEPTOR PI SUBUNIT	563598	GB	U95367		Breast lobularcarcinoma,
5954	SERINE/THREONINE KINASE MARK1	565000	GB	Z83868	Colon1,	
6027	RETINOIC ACID RECEPTOR BETA-2	589706	GB	Y00291	Colon1,	
6090	PYRUVATE DEHYDROGENASE KINASE ISOFORM 4	594120	GB	U54617		Colon1,
6223	RYANODINE RECEPTOR	628595	GB	J05200		
6236	GS3955	629446	GB	D87119	Colon1,	
6249	ORPHAN G PROTEIN-COUPLED RECEPTOR (RDC1)	645026	GB	U67784		
6303	PROTO-ONCOGENE TYROSINE-PROTEIN KINASE ABL	663987	GB	X16416	Colon1,	
6341	CANNABINOID RECEPTOR 2	683442	GB	X74328		

LifeSpan Master ID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
6342	FRAP-RELATED PROTEIN	683506	GB	U49844		
6354	TYROSINE-PROTEIN KINASE TEC	685344	GB	D29767		Breast lobularcarcinoma,
6371	PROTEIN KINASE CLK1	701594	GB	M59287	Liver,	Kidney 105, Kidney renalcellcarcinoma,
6422	MAP KINASE KINASE 4 (MKK4)	726147	GB	L36870	Liver,	
6460	CELLULAR PROTO-ONCOGENE (C-MER)	728657	GB	U08023	Colon1,	Kidney renalcellcarcinoma,
6565	GLUTAMATE RECEPTOR 5 PRECURSOR	758315	GB	X66118	breast,	
6591	VIP2 RECEPTOR	768352	GB	X95097		
6601	KIAA0135	770837	GB	D50925		breast,
6660	EPIDERMAL GROWTH FACTOR RECEPTOR PRECURSOR	60493	GB	X00588		Kidney 105,
6664	ALPHA2-C4-ADRENERGIC RECEPTOR	60664	GB	J03853	Colon1,	breast, Breast lobularcarcinoma,
6786	DUAL SPECIFICITY PROTEIN KINASE TTK	29739	GB	M86699		Kidney renalcellcarcinoma, Liver,
6787	RECEPTOR PROTEIN-TYROSINE KINASE (HEK11)	29543	GB	L36642		
6793	LIM DOMAIN KINASE 1	31097	GB	D26309	Colon1,	
6797	ACTIVATED P21CDC42HS KINASE (ACK)	33211	GB	L13738		breast,

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
6808	B LYMPHOCYTE SERINE/THREONINE PROTEIN KINASE	37234	GB	U07349		breast, Breast lobularcarcinoma,
6820	SERINE/THREONINE-PROTEIN KINASE PRP4H (PRP4H)	40240	GB	U48736		breast,
6831	RETINOIC ACID RECEPTOR RXR-GAMMA	41089	GB	U38480	Colon1,	Prostate,
6840	SERINE KINASE SRPK2	43108	GB	U88666	Liver,	
6842	SERINE/THREONINE-SPECIFIC PROTEIN KINASE MINIBRAIN	43033	GB	D86550		Breast lobularcarcinoma,
6902	GLYCOGEN SYNTHASE KINASE 3	155521	GB	L40027		breast,
6905	VASOACTIVE INTESTIAL PEPTIDE RECEPTOR	155943	GB	L13288		
6923	CELL ADHESION KINASE BETA (CAKBETA)	43541	GB	U43522		Kidney renalcellcarcinoma,
6953	PROTEIN KINASE NEK3	49897	GB	Z29067		
7003	PISSLRE	182347	GB	X78342	Colon1,	
7058	GLYCOGEN SYNTHASE KINASE 3	20087	GB	L40027		Liver,
7073	SERINE KINASE	246240	GB	U09564	Breast lobularcarcinoma, Kidney 510,	
7114	YSK1	267182	GB	D63780		Colon1,
7144	C-FMS PROTO-ONCOGENE	277866	GB	X03663		

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
7179	RETINOIC ACID RECEPTOR GAMMA-1	298678	GB	M38258	Colon1,	Breast lobularcarcinoma,
7264	DEATH-ASSOCIATED PROTEIN KINASE 1 (DAP)	341971	GB	X76104		Breast lobularcarcinoma,
7293	CASEIN KINASE I, GAMMA 2 ISOFORM	346031	GB	U89896	Kidney renalcellcarcinoma,	
7307	SIGNAL RECOGNITION PARTICLE	357421	GB	X67813		Colon1,
7337	RIBOSOMAL PROTEIN S6 KINASE 2 (RPS6KA2)	376362	GB	L07597		
7349	PROTEIN KINASE NEK2	415089	GB	Z29066		breast,
7360	CASEIN KINASE II, ALPHA CHAIN	417362	GB	M55265		
7373	P58/GTA (GALACTOSYLTRANSFERASE ASSOCIATED PROTEIN	429403	GB	M37712		breast,
7413	PROTEIN KINASE C, MU TYPE	491336	GB	X75756		breast,
7426	CDC7-RELATED KINASE	503569	GB	AF015592		breast,
7445	SERINE/THREONINE-PROTEIN KINASE PLK	511794	GB	X75932	Colon1,	
7454	RECEPTOR TYROSINE KINASE (DTK)	526868	GB	U18934	Colon1,	
7474	CASEIN KINASE II BETA SUBUNIT	548498	GB	M30448		Breast lobularcarcinoma,

LifeSpan MasterID	GeneName	Image CloneID	DataBase ID	Accession Number	3X higher in normal tissue than cancer	3X higher in cancerous tissue than normal
7481	LIPID-ACTIVATED, PROTEIN KINASE PRK2	550355	GB	U33052		Breast lobularcarcinoma,
7543	CYTOPLASMIC TYROSINE-PROTEIN KINASE BMX	624566	GB	X83107	Kidney renalcellcarcinoma,	Kidney 1010,
7544	SERINE/THREONINE PROTEIN KINASE MO15	624688	GB	Y13120	Breast lobularcarcinoma,	
7552	DIHYDROPRYRIDINE-SENSITIVE L-TYPE, SKELETAL MUSCLE	628206	GB	L33798	Colon1,	
7579	RECEPTOR PROTEIN-TYROSINE KINASE TKT PRECURSOR	668442	GB	X74764		Colon1,

WHAT IS CLAIMED IS:

1. A method for diagnosing cancer of a tissue of interest in a patient, comprising detecting the overexpression or the underexpression of a cancer-associated molecule in the tissue of interest in the patient, where overexpression or underexpression of such molecule is indicative of a cancer according to Table 1.

2. The method of claim 1 wherein underexpression of the cancer-associated molecule is indicative of cancer, and wherein the cancer-associated molecule is underexpressed in the patient.

3. The method of claim 1 wherein overexpression of the cancer-associated molecule is indicative of cancer, and wherein the cancer-associated molecule is overexpressed in the patient.

4. The method of claim 1 comprising detecting an mRNA encoding the cancer-associated molecule.

5. The method of claim 1 comprising detecting the cancer-associated molecule in an immunoassay.

6. The method of claim 1 wherein the tissue of interest is a tissue selected from the group consisting of liver, breast, prostate, kidney and colon.

7. A method for inhibiting cancer comprising introducing into a cell a cancer-associated molecule according to Table 1, wherein underexpression of the cancer-associated molecule is indicative of cancer.

8. The method of claim 7 wherein said cancer-associated molecule is a nucleic acid encoding a cancer-associated protein.

9. The method of claim 7 wherein said cancer-associated molecule is a protein.

10. A method for arresting cancer comprising inhibiting a cancer-associated molecule according to Table 1, wherein overexpression of the cancer-associated molecule is indicative of cancer.

11. The method of claim 10 wherein said cancer-associated molecule is inhibited using an antisense polynucleotide.

12. The method of claim 10 wherein said cancer-associated molecule is inhibited using an antibody that specifically binds to the cancer-associated protein.

13. A method for identifying a modulator of cancer development in a cell, the method comprising:

culturing the cell in the presence of the modulator to form a first cell culture;

contacting RNA or cDNA from the first cell culture with a probe which comprises a polynucleotide sequence associated with cancer development, wherein the polynucleotide sequence is selected from the group consisting of sequences set out in Table 1;

determining whether the amount of the probe which hybridizes to the RNA or cDNA from the first cell culture is increased or decreased relative to the amount of the probe which hybridizes to RNA or cDNA from a second cell culture grown in the absence of said modulator; and

detecting the presence of a decreased proliferative potential, transformation and malignancy in the first cell culture relative to the second cell culture.

14. The method of claim 13 wherein the first and second cell cultures are obtained from a liver, kidney, breast, colon or prostate cell.

15. A method for modulating cancer development in a patient in need thereof, the method comprising administering to the patient a compound that modulates the malignant transformation of a cell.

16. The method of claim 15, wherein the modulator increases or decreases the expression of a nucleic acid sequence set out in Table 1.